B79

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 October 2001 (04.10.2001)

PCT

(10) International Publication Number
WO 01/72962 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US01/09410

(22) International Filing Date: 23 March 2001 (23.03.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/191,929 24 March 2000 (24.03.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 60/191,929 (CIP)
Filed on 24 March 2000 (24.03.2000)

(71) Applicant and
(72) Inventor: **SAATCIOGLU, Fahri** [TR/US]; 26 Halsey Avenue, Wellesley, MA 02482 (US).

(74) Agent: **BIEKER-BRADY, Kristina**; Clark & Elbing LLP, 176 Federal Street, Boston, MA 02110-2214 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/72962 A2

(54) **Title:** NOVEL PROSTATE-SPECIFIC OR TESTIS-SPECIFIC NUCLEIC ACID MOLECULES, POLYPEPTIDES, AND DIAGNOSTIC AND THERAPEUTIC METHODS

(57) **Abstract:** The invention provides novel prostate-specific or testis-specific nucleic acid molecules, polypeptides, antibodies, and modulatory compounds for use in methods of diagnosing, treating, and preventing diseases and conditions of the prostate and testis, such as cancer.

NOVEL PROSTATE-SPECIFIC OR TESTIS-SPECIFIC NUCLEIC ACID MOLECULES, POLYPEPTIDES, AND DIAGNOSTIC AND THERAPEUTIC METHODS

Field of The Invention

The invention generally relates to the treatment of disorders associated with prostate and testis dysfunction and cell proliferation, and specifically relates to the identification and use of novel genes for diagnosis and treatment of such disorders.

Background of The Invention

Genitourinary disorders are often difficult to diagnose and treat effectively because they are present non-specifically. Two causes of genitourinary disorders are disorders of the prostate gland and the testis.

The prostate is a variable sized gland located in the male pelvis, and is made up of several different cell types, including epithelial cells and stromal cells. Prostate-associated disorders include prostate cancer, benign prostatic hyperplasia, and prostatitis. The male hormone testosterone and other androgen related hormones have major roles in the growth and function of the prostate. The testis is also subject to many defects, including developmental anomalies, inflammation, and cancer.

In men, prostate cancer is the most commonly diagnosed cancer and the second leading cause of cancer mortality following skin cancer. In the initial stages, prostate cancer is dependent on androgens for growth, and this dependence is the basis for androgen ablation therapy. In most cases, however, prostate cancer progresses to an androgen-independent phenotype for which there is no effective therapy available at present.

Currently, there is limited information regarding the molecular details of prostate cancer progression. Several independent approaches resulted in the identification of a few highly prostate-enriched genes that may have unique roles in this process. The first such gene discovered was Prostate Specific Antigen (PSA), the detection of which is currently used as a diagnostic tool and also as a marker for the progression of prostate cancer, albeit with significant limitations. More recently, several additional prostate-enriched genes were identified including prostate-specific membrane antigen (PSMA), prostate carcinoma tumor antigen 1 (PCTA-1), NKX3.1, prostate stem cell antigen (PSCA), DD3, and PCGEM1.

It would be beneficial to provide reagents useful for the diagnosis and therapy of disorders associated with the prostate and the testis, as well as other tissues.

Summary of the Invention

The invention provides, in general, a novel prostate-specific or testis-specific nucleic acid molecules, polypeptides, antibodies, and modulatory compounds for use in methods of diagnosing, treating, and preventing diseases and conditions of the prostate and testis, such as cancer.

In a first aspect the invention provides a substantially pure prostate-specific or testis-specific polypeptide, including a sequence substantially identical to the sequence of any of SEQ ID NOS: 14, 29, 32, 34, 36, 41, or 53. In a preferred embodiment of the first aspect, the substantially pure prostate-specific or testis-specific polypeptide includes the sequence of any of SEQ ID NOS: 14, 29, 32, 34, 36, 41, or 53. In another preferred embodiment, the invention provides an isolated nucleic acid molecule encoding a polypeptide of the first aspect, for example a nucleic acid

molecule including the sequence of any of SEQ ID NOS: 23, 28, 31, 33, 35, 40, or 52. Preferably, the polypeptide is derived from a mammal, e.g., a human.

In a second aspect, the invention provides an isolated prostate-specific or testis-specific nucleic acid molecule including a sequence substantially identical to SEQ ID NOS: 1-12, 22, 27, 30, and 51.

In a third aspect, the invention provides an isolated prostate-specific or testis-specific nucleic acid molecule consisting essentially of SEQ ID NOS: 15-21, 24-26, 42-50, and 54-70.

In preferred embodiments of some of the above aspects, the invention provides a vector, a cell, a cell including the vector, and a non-human transgenic animal including the isolated nucleic acid molecules.

In a fourth aspect, the invention provides an isolated nucleic acid molecule that hybridizes under high stringency conditions to the complement of any of the sequences set forth in SEQ ID NOS: 1-12, 15-28, 30, 31, 33, 35, 40, 42-50, 51, 52, or 54-70, where the isolated nucleic acid molecule encodes a prostate-specific or testis-specific polypeptide.

In a fifth aspect, the invention provides an isolated nucleic acid molecule, where the nucleic acid molecule includes a sequence that is antisense to the coding strand of any of the prostate-specific or testis-specific nucleic acid molecules set forth in SEQ ID NOS: 1-12, 15-28, 30, 31, 33, 35, 40, 42-50, 51, 52, or 54-70, or a fragment thereof.

In a sixth aspect, the invention provides a probe for analyzing a prostate-specific or testis-specific gene or homolog or fragment thereof, the probe having greater than 55% nucleotide sequence identity to a sequence encoding any of SEQ ID NOS: 1-12, 15-28, 30, 31, 33, 35, 40, 42-50, 51, 52, or 54-70, or fragment thereof, where the fragment includes at least six amino acids, and the probe hybridizes under high stringency conditions to at least a portion of a prostate-specific or testis-specific nucleic acid

molecule. In a preferred embodiment of this aspect, the probe has 100% complementarity to a nucleic acid molecule encoding any of SEQ ID NOS: 1-12, 15-28, 30, 31, 33, 35, 40, 42-50, 51, 52, or 54-70, or fragment thereof, where the fragment comprises at least six amino acids, and said probe hybridizes under high stringency conditions to at least a portion of a prostate-specific or testis-specific nucleic acid molecule.

In a seventh aspect, the invention provides an antibody that specifically binds to a prostate-specific or testis-specific polypeptide that includes an amino acid sequence that is substantially identical to the amino acid sequence of any of SEQ ID NOS: 14, 29, 32, 34, 36, 41, or 53.

In an eighth aspect, the invention provides a method of detecting a prostate-specific or testis-specific gene or fragment thereof in a cell, the method including contacting the nucleic acid molecule of any of SEQ ID NOS: 1-12, 15-28, 30, 31, 33, 35, 40, 42-50, 51, 52, or 54-70, or a fragment thereof, where the fragment is greater than about 18 nucleotides in length, with a preparation of genomic DNA from the cell, under high stringency hybridization conditions, and detecting DNA sequences having about 55% or greater nucleotide sequence identity to any of SEQ ID NOS: 1-12, 15-28, 30, 31, 33, 35, 40, 42-50, 51, 52, or 54-70, thus identifying a prostate-specific or testis-specific gene or fragment thereof. Nucleotides encoding the polypeptides of SEQ ID NOS: 38, 39, or 71-73 can also be used in an embodiment of this aspect. In a preferred embodiment of this aspect, the method includes detecting a neoplastic or cancer cell in a patient predisposed to or at risk for cancer, for example, for prostate cancer.

In a ninth aspect, the invention provides a method for identifying a test compound that modulates the expression or activity of a prostate-specific or testis-specific polypeptide, the method including contacting the prostate-specific or testis-specific polypeptide with the test compound, and determining the effect of the test compound on the prostate-specific or

testis-specific polypeptide expression or activity. In a preferred embodiments of this aspect, the prostate-specific or testis-specific polypeptide includes an amino acid sequence substantially identical to the amino acid sequence of SEQ ID NOS: 14, 29, 32, 34, 36, 38, 39, 41, 53, or 71-73, and fragments and analogs thereof.

In a tenth aspect, the invention provides a method of treating a mammal having a disorder of the prostate or testis, the method including administering to the mammal a therapeutically effective amount of a compound that modulates the activity or expression of a prostate-specific or testis-specific polypeptide, where the compound has a beneficial effect on the disorder in the mammal. In preferred embodiments of this aspect, the disorder is prostate cancer, the mammal is a human, or the prostate-specific or testis-specific polypeptide includes an amino acid sequence substantially identical to the amino acid sequence of SEQ ID NOS: 14, 29, 32, 34, 36, 38, 39, 41, 53, or 71-73, and fragments and analogs thereof.

In an eleventh aspect, the invention provides a pharmaceutical composition including at least one dose of a therapeutically effective amount of a prostate-specific or testis-specific polypeptide or fragment thereof, in a pharmaceutically acceptable carrier, the composition being formulated for the treatment of a disorder of the prostate or testis.

In a twelfth aspect, the invention provides a kit for the analysis of a prostate-specific or testis-specific nucleic acid molecule, the kit including a nucleic acid molecule probe for analyzing a prostate-specific or testis-specific nucleic acid molecule present in a test subject.

In a thirteenth aspect, the invention provides a kit for the analysis of a prostate-specific or testis-specific polypeptide, the kit including an antibody for analyzing a prostate-specific or testis-specific polypeptide present in a test subject.

As used herein, by "polypeptide," "protein," or "polypeptide fragment" is meant a chain of two or more amino acids, regardless of any post-translational modification (*e.g.*, glycosylation or phosphorylation), constituting all or part of a naturally or non-naturally occurring polypeptide. By "post-translational modification" is meant any change to a polypeptide or polypeptide fragment during or after synthesis. Post-translational modifications can be produced naturally (such as during synthesis within a cell) or generated artificially (such as by recombinant or chemical means). A protein can be made up of one or more polypeptides.

By "substantially pure polypeptide" or "substantially pure and isolated polypeptide" is meant a polypeptide (or a fragment thereof) that has been separated from components that naturally accompany it. Typically, the polypeptide is substantially pure when it is at least 60%, by weight, free from the proteins and naturally occurring organic molecules with which it is naturally associated. Preferably, the polypeptide is a prostate-specific or a testis-specific polypeptide that is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight, pure. A substantially pure prostate-specific or a testis-specific polypeptide may be obtained by standard techniques, for example, by extraction from a natural source (e.g., prostate or testis tissue or cell lines), by expression of a recombinant nucleic acid encoding a prostate-specific or a testis-specific polypeptide, or by chemically synthesizing the polypeptide. Purity can be measured by any appropriate method, e.g., by column chromatography, polyacrylamide gel electrophoresis, or HPLC analysis.

A protein or polypeptide is substantially free of naturally associated components when it is separated from those contaminants that accompany it in its natural state. Thus, a protein that is chemically synthesized or produced in a cellular system different from the cell from which it naturally originates will be substantially free from its naturally associated

determine identity. The BLAST program is publicly available from NCBI and other sources (*BLAST Manual*, Altschul, et al., NCBI NLM NIH Bethesda, MD 20894). Searches can be performed in URLs such as the following http://www.ncbi.nlm.nih.gov/BLAST/unfinishedgenome.html; or http://www.tigr.org/cgi-bin/BlastSearch/blast.cgi. These software programs match similar sequences by assigning degrees of homology to various substitutions, deletions, and other modifications. Conservative substitutions typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid, asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine.

A nucleic acid molecule or polypeptide is said to be "substantially identical" to a reference molecule if it exhibits, over its entire length, at least 50%, 60%, or 70%, preferably at least 80% or 90%, more preferably at least 95%, and most preferably at least 99% identity to the sequence of the reference molecule. For polypeptides, the length of comparison sequences is at least 16 amino acids, preferably at least 20 amino acids or at least 25 amino acids, more preferably at least 35 amino acids, and most preferably, the full-length polypeptide. For nucleic acid molecules, the length of comparison sequences is at least 50 nucleotides, preferably at least 60 nucleotides, more preferably at least 75 nucleotides or at least 110 nucleotides, and most preferably, the full-length nucleic acid molecule. Alternatively, or additionally, two nucleic acid sequences are "substantially identical" if they hybridize under high stringency conditions.

By "isolated nucleic acid molecule," "substantially pure nucleic acid molecule," or "substantially pure and isolated nucleic acid molecule" is meant a nucleic acid molecule (for example, DNA) that is free of the genes that, in the naturally occurring genome of the organism from which the nucleic acid molecule of the invention is derived, flank the nucleic acid. The term includes, for example, a recombinant DNA that is incorporated

into a vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote; or that exists as a separate molecule (e.g., a cDNA or a genomic or cDNA fragment produced by PCR or restriction endonuclease digestion) independent of other sequences. It also includes a recombinant DNA that is part of a hybrid gene encoding additional polypeptide sequence.

By "antisense," as used herein in reference to nucleic acid molecules, is meant a molecule having a nucleic acid sequence, regardless of length, that is complementary to at least 75 nucleotides, and preferably at least 100, 150, or 200 nucleotides, of the coding strand of a nucleic acid molecule encoding a prostate-specific or a testis-specific polypeptide, as described herein. An antisense molecule may also include regulatory sequences such as transcription enhancers, hormone responsive elements, ribosomal- and RNA polymerase binding sites, etc., which may be located upstream or downstream of the coding region, and may have a distance of several ten base pairs to several ten thousand base pairs. An antisense nucleic acid molecule can be, for example, capable of preferentially lowering the production or expression of a prostate-specific or a testis-specific polypeptide encoded by a prostate-specific or a testis-specific nucleic acid molecule.

By "prostate-specific" or "testis-specific" nucleic acid molecule is meant a nucleic acid molecule, such as a genomic DNA, cDNA, or RNA (*e.g.*, mRNA) molecule, having at least 50, 60, or 75%, more preferably at least 80, 85, or 95%, and most preferably at least 99% amino acid identity to the nucleic acid molecules described herein, for example, in Figures 4, 11, and 14. In addition, a nucleic acid molecule having at least 50, 60, or 75%, more preferably at least 80, 85, or 95%, and most preferably at least 99% nucleotide identity to a nucleotide sequence encoding amino acids 1-200 of STMP1 (SEQ ID NO: 14), preferably encoding amino acids 40-150

of STMP1, can be considered a prostate-specific or testis-specific nucleic acid molecule. Specifically excluded from this definition is STEAP (AF186249) (Hubert, R. S. et al., *Proc Natl Acad Sci U S A* 96, 14523-14528, 1999) and nucleic acid molecule sequences set forth in or encoding ESTs AF132025, AF177862, BAB23615, BAA91839, BAB15559, and NP_032190.

A preferred prostate-specific nucleic acid molecule may be preferentially expressed in prostate tissue at a level that is at least 5-fold higher, preferably at least 10-fold higher, more preferably at least 15-fold higher, and most preferably at least 20-fold higher than the level of the same nucleic acid molecule in at least one non-prostate tissue, preferably in all other non-prostate tissues. A prostate-specific nucleic acid molecule can also be expressed at high levels in a non-prostate tissue although, generally, the level of expression will be the highest in the prostate. Occasionally, as described herein, a prostate-specific nucleic acid molecule will be expressed at higher levels in non-prostate tissue (e.g., placenta, lung, or liver) than in the prostate.

A preferred testis-specific nucleic acid molecule may be preferentially expressed in testis tissue at a level that is at least 5-fold higher, preferably at least 10-fold higher, more preferably at least 15-fold higher, and most preferably at least 20-fold higher than the level of the same nucleic acid molecule in at least one non- testis tissue, preferably in all other non- testis tissues. A testis -specific nucleic acid molecule can also be expressed at high levels in a non- testis tissue although, generally, the level of expression will be the highest in the testis. Occasionally, as described herein, a testis -specific nucleic acid molecule will be expressed at higher levels in non- testis tissue (e.g., placenta, lung, or liver) than in the testis.

By "prostate-specific" or a "testis-specific" polypeptide or "prostate-specific" or a "testis-specific" protein is meant a polypeptide that is encoded by a prostate-specific or a testis-specific nucleic acid molecule. A prostate-specific or testis-specific polypeptide may also be defined as a polypeptide having at least 50, 60, or 75%, more preferably at least 80, 85, or 95%, and most preferably at least 99% amino acid identity to the polypeptides described herein, for example, in Figures 4, 11, and 14. Specifically excluded from this definition is STEAP (AF186249) (Hubert, R. S. et al., *Proc Natl Acad Sci U S A* 96, 14523-14528, 1999) and polypeptide sequences set forth in or encoded by ESTs AF132025, AF177862, BAB23615, BAA91839, BAB15559, and NP_032190. In addition, a polypeptide having at least 50, 60, or 75%, more preferably at least 80, 85, or 95%, and most preferably at least 99% amino acid identity to amino acids 1-200 of STMP1 (SEQ ID NO: 14), preferably amino acids 40-150 of STMP1, can be considered a prostate-specific or testis-specific polypeptide.

A preferred prostate-specific polypeptide is preferentially expressed in prostate tissue at a level that is at least 5-fold higher, preferably at least 10-fold higher, more preferably at least 15-fold higher, and most preferably at least 20-fold higher than the level of the same polypeptide in at least one non-prostate tissue, preferably in all other non-prostate tissues. A prostate-specific polypeptide can also be expressed at high levels in a non-prostate tissue although, generally, the level of expression will be the highest in the prostate. Occasionally, as described herein, a prostate-specific polypeptide will be expressed at higher levels in non-prostate (e.g., placenta, lung, liver) than in the prostate.

A preferred testis-specific polypeptide is preferentially expressed in testis tissue at a level that is at least 5-fold higher, preferably at least 10-fold higher, more preferably at least 15-fold higher, and most preferably at

least 20-fold higher than the level of the same polypeptide in at least one non- testis tissue, preferably in all other non-testis tissues. A testis-specific polypeptide can also be expressed at high levels in a non- testis tissue although, generally, the level of expression will be the highest in the testis. Occasionally, as described herein, a testis-specific polypeptide will be expressed at higher levels in non- testis (e.g., placenta, lung, liver) than in the testis.

The term prostate-specific or testis-specific polypeptide includes homologs, analogs, fragments, and isoforms, e.g., alternatively spliced isoforms, of the sequences described herein. By "biologically active fragment" is meant a polypeptide fragment of a prostate-specific or testis-specific polypeptide that exhibits, for example, extracellular trafficking, cell signaling, or other properties that are at least 30%, preferably at least 50%, more preferably at least 75%, and most preferably at least 100%, compared with the properties of a full length prostate-specific or testis-specific polypeptide. By "analog" is meant any substitution, addition, or deletion in the amino acid sequence of a prostate-specific or testis-specific polypeptide that exhibits properties that are at least 30%, preferably at least 50%, more preferably at least 75%, and most preferably at least 100%, compared with the extracellular trafficking or cell signaling properties of the polypeptide from which it is derived. Fragments, homologs, and analogs can be generated using standard techniques, for example, solid phase peptide synthesis or polymerase chain reaction. For example, point mutations may arise at any position of the sequence from an apurinic, apyrimidinic, or otherwise structurally impaired site within the cDNA. Alternatively, point mutations may be introduced by random or site-directed mutagenesis procedures (*e.g.*, oligonucleotide assisted or by error prone PCR). Likewise, deletions and/or insertions may be introduced into the sequences, and preferred insertions comprise 5'- and/or 3'-fusions with a

polynucleotide that encodes a reporter moiety or an affinity moiety. Other preferred insertions comprise a nucleic acid that further includes functional elements such as a promoter, enhancer, hormone responsive element, origin of replication, transcription and translation initiation sites, etc. It should be appreciated that where insertions with one or more functional elements are present, the resulting nucleic acid may be linear or circular (*e.g.*, transcription or expression cassettes, plasmids, etc.).

For use in the methods of the invention, the terms "prostate-specific" or "testis-specific" polypeptide further include the polypeptide sequences set forth in or encoded by ESTs AF132025, AF177862, BAB23615, BAA91839, BAB15559, and NP_032190, but does not include STEAP, and a prostate-specific or testis-specific nucleic acid molecule includes the nucleotide sequences set forth in or encoding ESTs AF132025, AF177862, BAB23615, BAA91839, BAB15559, and NP_032190, but does not include STEAP.

By "prostate-specific or a testis-specific gene or homolog or fragment thereof" is meant a gene, or homolog of a gene, that encodes a prostate-specific or testis-specific polypeptide.

By "specifically binds" is meant a compound, e.g., an antibody, that recognizes and binds a protein or polypeptide, for example, a prostate-specific or a testis-specific polypeptide, and that when detectably labeled can be competed away for binding to that protein or polypeptide by an excess of compound that is not detectably labeled. A compound that non-specifically binds is not competed away by excess detectably labeled compound. A preferred antibody binds to any prostate-specific or a testis-specific polypeptide sequence that is substantially identical to any of the polypeptide sequences set forth in Figures 4, 11, and 14, or encoded by any of the nucleotide sequences set forth in Figures 3, 4, 11, and 14, or portions thereof.

By a "compound," "test compound," or "candidate compound" is meant a molecule, be it naturally-occurring or artificially-derived, and includes, for example, peptides, proteins, synthetic organic molecules, naturally-occurring organic molecules, nucleic acid molecules, and components thereof.

By "high stringency conditions" is meant conditions that allow hybridization comparable with the hybridization that occurs using a DNA probe of at least 500 nucleotides in length, in a buffer containing 0.5 M $NaHPO_4$, pH 7.2, 7% SDS, 1 mM EDTA, and 1% BSA (fraction V), at a temperature of 65°C, or a buffer containing 48% formamide, 4.8X SSC, 0.2 M Tris-Cl, pH 7.6, 1X Denhardt's solution, 10% dextran sulfate, and 0.1% SDS, at a temperature of 42°C (these are typical conditions for high stringency Northern or Southern hybridizations). High stringency hybridization is also relied upon for the success of numerous techniques routinely performed by molecular biologists, such as high stringency PCR, DNA sequencing, single strand conformational polymorphism analysis, and *in situ* hybridization. In contrast to Northern and Southern hybridizations, these techniques are usually performed with relatively short probes (e.g., usually 16 nucleotides or longer for PCR or sequencing, and 40 nucleotides or longer for *in situ* hybridization). The high stringency conditions used in these techniques are well known to those skilled in the art of molecular biology, and may be found, for example, in Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY, 1998, hereby incorporated by reference.

By "probe" or "primer" is meant a single-stranded DNA or RNA molecule of defined sequence that can base pair to a second DNA or RNA molecule that contains a complementary sequence ("target"). The stability of the resulting hybrid depends upon the extent of the base pairing that occurs. This stability is affected by parameters such as the degree of

complementarity between the probe and target molecule, and the degree of stringency of the hybridization conditions. The degree of hybridization stringency is affected by parameters such as the temperature, salt concentration, and concentration of organic molecules, such as formamide, and is determined by methods that are well known to those skilled in the art. Probes or primers specific for prostate-specific or a testis-specific nucleic acid molecules, preferably, have greater than 45% sequence identity, more preferably at least 55-75% sequence identity, still more preferably at least 75-85% sequence identity, yet more preferably at least 85-99% sequence identity, and most preferably 100% sequence identity to the nucleic acid sequences encoding the amino acid sequences described herein. Probes can be detectably-labeled, either radioactively or non-radioactively, by methods that are well-known to those skilled in the art. Probes can be used for methods involving nucleic acid hybridization, such as nucleic acid sequencing, nucleic acid amplification by the polymerase chain reaction, single stranded conformational polymorphism (SSCP) analysis, restriction fragment polymorphism (RFLP) analysis, Southern hybridization, northern hybridization, *in situ* hybridization, electrophoretic mobility shift assay (EMSA), and other methods that are well known to those skilled in the art.

A molecule, *e.g.*, an oligonucleotide probe or primer, a gene or fragment thereof, a cDNA molecule, a polypeptide, or an antibody, can be said to be "detectably-labeled" if it is marked in such a way that its presence can be directly identified in a sample. Methods for detectably-labeling molecules are well known in the art and include, without limitation, radioactive labeling (*e.g.*, with an isotope, such as ^{32}P or ^{35}S) and nonradioactive labeling (*e.g.*, with a fluorescent label, such as fluorescein, or by generating a construct containing green fluorescent protein (GFP)).

By "transgenic" is meant any cell that includes a DNA sequence or transgene that is inserted by artifice into a cell and becomes part of the genome of the organism that develops from that cell. As used herein, the transgenic organisms are generally transgenic mammals (e.g., mice, rats, and goats) and the DNA (transgene) is inserted by artifice into the nuclear genome. By "transgene" is meant any piece of DNA that is inserted by artifice into a cell, and becomes part of the genome of the organism that develops from that cell. Such a transgene may include a gene that is partly or entirely heterologous (i.e., foreign) to the transgenic organism, or may represent a gene homologous to an endogenous gene of the organism. By "knockout mutation" is meant an artificially induced alteration in the nucleic acid sequence (created via recombinant DNA technology or deliberate exposure to a mutagen) that reduces the biological activity of the polypeptide normally encoded therefrom by at least 80% relative to the unmutated gene. The mutation may, without limitation, be an insertion, deletion, frameshift mutation, or a missense mutation. The knockout mutation can be in a cell *ex vivo* (e.g., a tissue culture cell or a primary cell) or *in vivo*. A "knockout animal" is a mammal, preferably, a mouse, containing a knockout mutation as defined above.

By "sample" is meant a tissue biopsy, cells, blood, serum, urine, stool, or other specimen obtained from a patient or test subject. The sample is analyzed to detect a mutation in a gene encoding a prostate-specific or a testis-specific polypeptide, or expression levels of a gene encoding a prostate-specific or a testis-specific polypeptide, as for example, an indication of the progression of cancer, by methods that are known in the art or described herein. For example, methods such as sequencing, single-strand conformational polymorphism (SSCP) analysis, or restriction fragment length polymorphism (RFLP) analysis of PCR products derived from a patient sample may be used to detect a mutation in a gene encoding

a prostate-specific or a testis-specific polypeptide; ELISA may be used to measure levels of a prostate-specific or a testis-specific polypeptide; and PCR may be used to measure the level of nucleic acids encoding a prostate-specific or a testis-specific polypeptide.

By "pharmaceutically acceptable carrier" is meant a carrier that is physiologically acceptable to the treated mammal while retaining the therapeutic properties of the compound with which it is administered. One exemplary pharmaceutically acceptable carrier is physiological saline solution. Other physiologically acceptable carriers and their formulations are known to one skilled in the art and described, for example, in *Remington: The Science and Practice of Pharmacy*, (19^{th} edition), ed. A. Gennaro, 1995, Mack Publishing Company, Easton, PA.

"Therapeutically effective amount" as used herein in reference to dosage of a medication, refers to the administration of a specific amount of a pharmacologically active agent (*e.g.*, a prostate-specific or a testis-specific polypeptide, nucleic acid molecule, or modulatory compound) tailored to each individual patient manifesting symptoms characteristic of a specific disorder. For example, a patient receiving the treatment of the present invention might have prostate cancer. A person skilled in the art will recognize that the optimal dose of a pharmaceutical agent to be administered will vary from one individual to another. Dosage in individual patients should take into account the patients height, weight, rate of absorption and metabolism of the medication in question, the stage of the disorder to be treated, and what other pharmacological agents are administered concurrently.

By "treating " or "treatment" is meant the medical management of a patient with the intent to cure, ameliorate, stabilize, or prevent a disease, pathological condition, or disorder. This term includes active treatment, that is, treatment directed specifically toward the improvement or

associated with the cure of a disease, pathological condition, or disorder, and also includes causal treatment, that is, treatment directed toward removal of the cause of the associated disease, pathological condition, or disorder. In addition, this term includes palliative treatment, that is, treatment designed for the relief of symptoms rather than the curing of the disease, pathological condition, or disorder; preventative treatment, that is, treatment directed to minimizing or partially or completely inhibiting the development of the associated disease, pathological condition, or disorder; and supportive treatment, that is, treatment employed to supplement another specific therapy directed toward the improvement of the associated disease, pathological condition, or disorder. The phrase "treatment" also includes symptomatic treatment, that is, treatment directed toward constitutional symptoms of the associated disease, pathological condition, or disorder.

By "disorder of the prostate or testis" is meant a disturbance of function and/or structure of the prostate or testis in a living organism, resulting from an external source, a genetic predisposition, a physical or chemical trauma, or a combination of the above. Such disorders include the proliferation of prostate or testicular cells. By "cell proliferation" is meant the growth or reproduction of similar cells, and the invention provides reagents for inhibiting proliferation and stimulating proliferation. By "inhibiting proliferation" is meant the decrease in the number of similar cells by at least 10%, more preferably by at least 20%, and most preferably by at least 50%. By "stimulating proliferation" is meant an increase in the number of similar cells by at least 10%, more preferably by at least 20%, and most preferably by at least 50%.

The reagents described herein, for example, vectors expressing antisense, antagonists, or inhibitors of prostate-specific or testis-specific polypeptides or nucleic acid molecules may be used, for example, to

suppress the excessive proliferation of prostate or testicular cells. Blocking prostate-specific or testis-specific polypeptide or nucleic acid molecule expression or activity in prostate or testicular cells can alter molecular pathways within cancerous cells and thus trigger apoptosis, i.e., the process of cell death where a dying cell displays a set of well-characterized biochemical hallmarks which include cytolemmal blebbing, cell soma shrinkage, chromatin condensation, and DNA laddering.

Disorders of the prostate or testis include prostate cancer, benign prostatic hyperplasia, acute prostatitis, testicular cancer, developmental defects of the prostate or testis (such as cryptorchidism or undescended testis, and retractile, ascending, or vanished testis).

By "proliferative disease" is meant a disease that is caused by or results in inappropriately high levels of cell division, inappropriately low levels of apoptosis, or both. For example, cancers such as prostate cancer, testicular cancer, lymphoma, leukemia, melanoma, ovarian cancer, breast cancer, pancreatic cancer, liver cancer, and lung cancer are all examples of proliferative disease.

By "modulate" or "modulating" is meant changing, either by decrease or increase, the expression or biological activity of a prostate-specific or testis-specific nucleic acid molecule or polypeptide, as described herein. It will be appreciated that the degree of modulation provided by a modulating compound in a given assay will vary, but that one skilled in the art can determine the statistically significant change in the level of biological activity that identifies a compound that modulates a prostate-specific or testis-specific nucleic acid molecule or polypeptide.

The invention provides several advantages. For example, it provides methods and reagents that can be used in the diagnosis and treatment of prostate and testis associated diseases, as well as other disorders and conditions that are sensitive to the bioactivities of the reagents (e.g.,

polypeptides, nucleic acid molecules, antibodies) described herein. Since the prostate-specific or testis-specific polypeptides of the invention have been found to be highly expressed in the prostate and testis, these polypeptides can also be used in screens for therapeutics to treat disorders associated with the prostate and testis. These polypeptides are also expressed in other tissues, and can be used as therapeutics and diagnostics for cell proliferative disorders.

Other features and advantages of the invention will be apparent from the detailed description of the invention, the drawings, and the claims.

Brief Description of The Drawings

Figure 1 shows an exemplary reverse northern analysis of several clones from a prostate specific cDNA library.

Figure 2 shows an exemplary multiple tissue northern blot.

Figure 3 is a table showing the nucleotide sequences of twelve clones (SEQ ID NOs: 1-12) isolated from prostate tissue and LNCaP cells .

Figure 4A is a schematic diagram showing the *STMP1* gene structure.

Figure 4B shows the nucleotide sequence, including the intron junction sequences (SEQ ID NO: 13), and predicted amino acid sequence (SEQ ID NO: 14) of STMP1.

Figure 4C shows the nucleotide sequences of the exons and 3' UTR of *STMP1* (SEQ ID NOs: 15-21).

Figure 4D shows the nucleotide sequence of the ORF of *STMP1* (SEQ ID NO: 22).

Figure 4E shows the shows the cDNA sequence (SEQ ID NO: 23), and predicted amino acid sequence (SEQ ID NO:14) of STMP1.

Figure 4F shows the nucleotide sequences of the exons and 3' UTR of *STMP1* ORF2 (SEQ ID NOs: 17-20 and 24-26).

Figure 4G shows the nucleotide sequence of the ORF of *STMP1* ORF2 (SEQ ID NO: 27).

Figure 4H shows the cDNA sequence (SEQ ID NO: 28), and predicted amino acid sequence (SEQ ID NO: 29) of STMP1 ORF2.

Figure 4I shows the nucleotide sequences of the exons and 3' UTR of *STMP1* ORF3 (SEQ ID NOs: 17-19 and 24-26).

Figure 4J shows the nucleotide sequence of the ORF of *STMP1* ORF3 (SEQ ID NO: 30).

Figure 4K shows the cDNA sequence (SEQ ID NO: 31), and predicted amino acid sequence (SEQ ID NO: 32) of STMP1 ORF3.

Figure 4L shows the cDNA sequence (SEQ ID NO: 33), and predicted amino acid sequence (SEQ ID NO:34) of STMP2.

Figure 4M shows the cDNA sequence (SEQ ID NO: 35), and predicted amino acid sequence (SEQ ID NO: 36) of STMP3.

Figure 5 shows a sequence alignment of STMP1 (SEQ ID NO: 14), with STEAP (SEQ ID NO: 37, Accession No. AF186249), and two ESTs (Accession No. BAA91839 and Accession No. BAB15559; SEQ ID NOs: 38 and 39, respectively).

Figure 6A shows a multiple tissue Northern blot probed with *STMP1* or *G3PDH* cDNA.

Figure 6B is a Northern blot probed with *STMP1* and *PSA* in the androgen-responsive prostate cancer cell line LNCaP and in the CWR22 human prostate cancer xenograft model.

Figure 6C is a Northern blot probed with *STMP1* and *NKX3A* in LNCaP, PC-3, and DU-145 cell lines and in the CWR22R human prostate cancer xenograft model.

Figure 7A shows fluorescence microscopy images of COS-1 cells transiently transfected with GFP-STMP1.

Figure 7B shows fluorescence microscopy images of COS-1 cells transiently transfected with GFP-STMP1 and labeled with antibodies against Golgi markers.

Figure 8 shows fluorescence microscopy images of COS-1 cells transiently transfected with GFP-STMP1 and observed by live-cell confocal microscopy.

Figure 9 shows fluorescence microscopy images of COS-1 cells transiently transfected with GFP-STMP1 and labeled with an antibody against an early endosomal marker.

Figure 10 is a schematic diagram showing the *SSH9* gene structure and two mRNA species transcribed from the *SSH9* gene.

Figure 11A shows the cDNA (SEQ ID NO: 40) and predicted amino acid sequence (SEQ ID NO: 41) for *SSH9*.

Figure 11B shows the predicted promoter sequence for *SSH9* (SEQ ID NO: 42).

Figure 11C shows the predicted intron-exon boundaries for *SSH9* (SEQ ID NOs: 43-50).

Figure 12A is a Northern blot probed with *SSH9* in the androgen-responsive prostate cancer cell line LNCaP cells and in the CWR22 human prostate cancer xenograft model.

Figure 12B is a Northern blot probed with *SSH9* in LNCaP, PC-3, and DU-145 cell lines, and CWR22R human prostate cancer xenograft model.

Figure 12C is a multiple tissue Northern blot probed with *SSH9* or *GAPDH* cDNA.

Figure 13 is a schematic diagram showing the *PSL22* gene structure.

Figure 14A shows the nucleotide sequence of the ORF of *PSL22* (SEQ ID NO: 51).

Figure 14B shows the cDNA sequence (SEQ ID NO: 52), and predicted amino acid sequence (SEQ ID NO: 53) of PSL22.

Figure 14C shows the nucleotide sequences of the TATA promoter and transcription start site, exons, and 5' and 3' UTRs of *PSL22* (SEQ ID NOs: 54-70).

Figure 15 shows a sequence alignment of PSL22 (RhoBP) (SEQ ID NO: 53), with ESTs NP032190 (mRhoph), AF132025 (dRhoph), and BAB23615 (SEQ ID Nos:71-73).

Figure 16A is a Northern blot probed with *PSL22* in LNCaP, PC-3, and DU-145 cell lines, and in the CWR22R human prostate cancer xenograft model.

Figure 16B is a multiple tissue Northern blot probed with *PSL22* cDNA.

Detailed Description of the Invention

The basic biology of the normal prostate and testis, as well as prostate and testicular cancer initiation and progression is still poorly understood. It is therefore necessary to delineate the molecular events that are at the basis of these processes. To achieve this goal, we have identified, cloned, and characterized highly prostate- and testis-enriched genes whose gene products have important roles for both the normal physiology and the pathophysiology of the prostate and the testis. These gene products also have important roles in other disorders, for example, heart, brain, liver, pancreas, kidney, and colon, which are the tissues where variable low expression, and occasionally, very high expression of specific gene products, can be detected by Northern analysis.

The invention provides prostate-specific or testis-specific polypeptides and nucleic acid molecules (see below), and diagnostic and therapeutic methods employing these polypeptides and nucleic acid

molecules. The invention also provides methods for identifying compounds that modulate the biological activities of prostate-specific or testis-specific polypeptides and nucleic acid molecules, and therapeutic methods employing these compounds. The diagnostic, therapeutic, and screening methods of the invention are first described, followed by general approaches that can be used in carrying out these methods. Finally, experimental results supporting the methods of the invention are described.

Bioassays

Prostate-specific and testis-specific polypeptides are expressed in the prostate and testis, and also in other tissues such as kidney, pancreas, liver, lung, and colon. The expression patterns of prostate-specific and testis-specific polypeptides in specific cells and tissues are used to identify cellular targets of prostate-specific and testis-specific polypeptide actions, and to identify bioactivities that are relevant to specific prostate- and testis-related diseases, such as prostate cancer, testicular cancer, benign prostatic hyperplasia, acute prostatitis, and developmental testis defects.

Therapeutic and diagnostic utilities for prostate-specific and testis-specific polypeptides are identified by, for example, conducting bioassays *in vitro*. Culture systems that reflect prostate-specific and testis-specific polypeptide expression patterns, along with the distribution of particular receptors, such as the androgen receptor, are selected. For example, LNCaP cells express androgen receptors, and respond to one or more isoforms of prostate-specific and testis-specific polypeptides in a variety of bioassays. The activities of prostate-specific and testis-specific polypeptides (*e.g.*, STMP1, SSH9, PSL22) are compared, using sister cultures, in various dose-response assays, including but not limited to, inhibition of proliferation, apoptosis, signaling events (e.g. changes in kinase activity), changes in transcription factor activity (such as that of the androgen receptor), intracellular trafficking, or cell signaling. The relative

potencies of the prostate-specific and testis-specific polypeptides are determined on the basis of, for example, protein concentration.

Diagnostic Methods Employing Prostate-Specific Or Testis-Specific Nucleic Acid Molecules, Polypeptides, and Antibodies

Prostate-specific or testis-specific nucleic acid molecules, polypeptides, and antibodies are used in methods to diagnose or monitor a variety of diseases and conditions, including those involving mutations in, or inappropriate expression of, prostate-specific or testis-specific genes. Prostate-specific or testis-specific expression has been documented in a variety of tissues, as discussed above. Thus, detection of abnormalities in prostate-specific or testis-specific genes or their expression is used in methods to diagnose, or to monitor treatment or development of diseases of these tissues.

The diagnostic methods of the invention are used, for example, with patients that have a prostate-related or testis-related disease, for example, prostate or testicular cancer, in an effort to determine its etiology, and thus, to facilitate selection of an appropriate course of treatment. The diagnostic methods are also used with patients that have not yet developed a prostate-related or testis-related disease, but who may be at risk of developing such a disease, or with patients that are at an early stage of developing such a disease. Many prostate-related or testis-related diseases occur during development, and thus, the diagnostic methods of the invention are also carried out on a fetus or embryo during development. Also, the diagnostic methods of the invention are used in prenatal genetic screening, for example, to identify parents who may be carriers of a recessive prostate-related or testis-related mutation.

Prostate-specific or testis-specific abnormalities that are detected using the diagnostic methods of the invention include those characterized

by, for example, (i) abnormal prostate-specific or testis-specific polypeptides, (ii) prostate-specific or testis-specific genes containing mutations that result in the production of such polypeptides, and (iii) mutations that result in production of abnormal amounts of prostate-specific or testis-specific polypeptides.

Levels of prostate-specific or testis-specific expression in a patient sample are determined by using any of a number of standard techniques that are well known in the art. For example, prostate-specific or testis-specific expression in a biological sample (*e.g.*, a blood, prostate or testis tissue sample, or amniotic fluid) from a patient is monitored by standard northern blot analysis or by quantitative PCR (see, *e.g.*, Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY, 1998; *PCR Technology: Principles and Applications for DNA Amplification*, H.A. Ehrlich, Ed., Stockton Press, NY; Yap *et al. Nucl. Acids. Res.* 19:4294, 1991).

A biological sample obtained from a patient can be analyzed for one or more mutations in prostate-specific or testis-specific nucleic acid molecules using a mismatch detection approach. Generally, this approach involves PCR amplification of nucleic acid molecules from a patient sample, followed by identification of a mutation (*i.e.*, a mismatch) by detection of altered hybridization, aberrant electrophoretic gel migration, binding, or cleavage mediated by mismatch binding proteins, or by direct nucleic acid molecule sequencing. Any of these techniques can be used to facilitate detection of mutant prostate-specific or testis-specific genes, and each is well known in the art. Examples of these techniques are described, for example, by Orita *et al.* (*Proc. Natl. Acad. Sci. USA* 86:2766-2770, 1989) and Sheffield *et al.* (*Proc. Natl. Acad. Sci. USA* 86:232-236, 1989).

Mismatch detection assays also provide an opportunity to diagnose a prostate-specific or testis-specific gene-mediated predisposition to a disease

before the onset of symptoms. For example, a patient heterozygous for a prostate-specific or testis-specific mutation that suppresses normal prostate-specific or testis-specific biological activity or expression may show no clinical symptoms of a prostate-specific or testis-specific gene-related disease, and yet possess a higher than normal probability of developing a prostate or testicular disease. Given such a diagnosis, patients can take precautions to minimize their exposure their exposure to adverse environmental factors and to carefully monitor their medical condition (for example, through frequent physical examinations). As mentioned above, this type of diagnostic approach can also be used to detect prostate-specific or testis-specific mutations in prenatal screens.

The prostate-specific or testis-specific diagnostic assays described above can be carried out using any biological sample (for example, a blood, prostate, or testis tissue sample, or amniotic fluid) in which a prostate-specific or testis-specific polypeptide or nucleic acid molecule is normally expressed. A mutant prostate-specific or testis-specific gene can also be identified using these sources as test samples. Alternatively, a prostate-specific or testis-specific mutation, as part of a diagnosis for predisposition to a prostate-specific or testis-specific gene-associated disease, can be tested for using a DNA sample from any cell, for example, by mismatch detection techniques. Preferably, the DNA sample is subjected to PCR amplification prior to analysis.

In yet another diagnostic approach of the invention, an immunoassay is used to detect or monitor prostate-specific or testis-specific protein expression in a biological sample. Anti-prostate-specific or testis-specific--polypeptide polyclonal or monoclonal antibodies (as described below) can be used in any standard immunoassay format (*e.g.*, ELISA, Western blot, or RIA; see, *e.g.*, Ausubel *et al.*, *supra*) to measure prostate-specific or testis-specific polypeptide levels. These levels are compared to wild-type

prostate-specific or testis-specific levels. For example, an increase in prostate-specific or testis-specific polypeptide production may be indicative of a condition or a predisposition to a condition involving overexpression of prostate-specific or testis-specific polypeptide biological activity, such as late stage prostate cancer.

Immunohistochemical techniques can also be utilized for prostate-specific or testis-specific polypeptide detection. For example, a tissue sample can be obtained from a patient, sectioned, and stained for the presence of prostate-specific or testis-specific polypeptide using an anti-prostate-specific or testis-specific antibody (see below) and any standard detection system (*e.g.*, one that includes a secondary antibody conjugated to horseradish peroxidase). General guidance regarding such techniques can be found in, *e.g.*, Bancroft *et al.*, *Theory and Practice of Histological Techniques*, Churchill Livingstone, 1982, and Ausubel *et al.*, *supra*.

In a preferred example, a combined diagnostic method can be employed that includes an evaluation of prostate-specific or testis-specific protein production (for example, by immunological techniques or the protein truncation test (Hogerrorst *et al.*, *Nature Genetics* 10:208-212, 1995), and a nucleic acid molecule-based detection technique designed to identify more subtle prostate-specific or testis-specific mutations (for example, point mutations). As described above, a number of mismatch detection assays are available to those skilled in the art, and any preferred technique can be used. Mutations in prostate-specific or testis-specific genes can be detected that either result in loss or gain of prostate-specific or testis-specific polypeptide or nucleic acid molecule expression or loss or gain of normal prostate-specific or testis-specific polypeptide or nucleic acid molecule biological activity.

Prostate-specific or testis-specific polypeptides or nucleic acid molecules can be used to correlate the course of prostate cancer to a marker

other than PSA, to monitor the course of an anticancer therapy, or to detect a neoplastic cell in a system. For example, a predetermined quantity of an RNA encoding a prostate-specific or testis-specific polypeptide is correlated with the presence of a neoplastic cell, for example, from a biopsy. The total RNA is extracted from the biopsy specimen, and a real time quantitative rt-PCR employing individual reactions with primer pairs specific to prostate-specific or testis-specific sequences is performed in parallel with a biopsy specimen known to be free of cancer cells. Biopsy specimens are determined to have a cancer cell, where the detected prostate-specific or testis-specific mRNA quantity is at least 5 times higher than in the control specimen. An exemplary extraction of total RNA utilizes the Quiagen BioRobot kit in conjunction with the BioRobot 9600 system, and the real time rtPCR is performed in a Perkin Elmer ABI Prism 7700.

In alternative aspects of the inventive subject matter, the method of detecting a neoplastic cell need not be limited to biopsy tissues from prostate or testis tissue, but may employ various alternative tissues, including lymphoma tumor cells, and various solid tumor cells, so long as such tumor cells overproduce mRNA of prostate-specific or testis-specific polypeptides. Appropriate alternative tumor cells can readily be identified by the above described method. Likewise, the system need not be restricted to a mammal, but may also include cell-, and tissue cultures grown *in vitro*, and tumor cells and specimens from animals other than mammals. For example, tumor cell and tissue grown *in vitro* may advantageously be utilized to investigate drug action on such cells, and sequences encoding prostate-specific or testis-specific polypeptides may conveniently be employed as tumor marker. Alternatively, body fluids (*e.g.*, serum, saliva, etc.) that may or may not contain tumor cells are also contemplated a suitable substrate for the method presented herein, so long as they contain

to at least some extent mRNA encoding a prostate-specific or testis-specific polypeptide.

In still other aspects of contemplated methods, the polypeptide quantity need not necessarily be limited to at least 5 times more than the control specimen in order to establish that the tissue has a cancer cell. For example, where the concentration of the polypeptide is hormone dependent, amounts between 3-8 fold and more may be appropriate. In contrast, where the concentration of cancer cells in the biopsy specimen is relatively low, amounts of less than 5-fold, including 1.5 to 4.9-fold and less are contemplated.

The detection process may include fluorescence detection, luminescence detection, scintigraphy, autoradiography, and formation of a dye. For example, for microscopic analysis of biopsy specimens, luciferase labeled probes are particularly advantageous in conjunction with a luminescence substrate (*e.g.*, luciferin). Luminescence quantification may then be performed utilizing a CCD-camera and image analysis system. Similarly, radioactivity may be detected via autoradiographic or scintigraphic procedures on a tissue section, in a fluid or on a solid support. Where the probe is a natural or synthetic ligand of a prostate-specific or testis-specific polypeptide, the ligand may include molecules with a chemical modification that increase the affinity to the polypeptide and/or induce irreversible binding to the polypeptide. For example, transition state analogs or suicide inhibitors for a particular reaction catalyzed by the polypeptide are especially contemplated. Labeling of antibodies, antibody fragments, small molecules, and binding of the labeled entity is a technique that is well known in the art, and all known methods are generally suitable for use in conjunction with methods contemplated herein. Furthermore, the probe need not be limited to a fluorescein labeled antibody, and alternative probes include antibody fragments (*e.g.*, Fab, Fab', scFab, etc.).

Still further contemplated variations include substitution of one or more atoms or chemical groups in the sequence with a radioactive atom or group. For example, where cDNAs are employed as a hybridization-specific probes, a fluorophor or enzyme (*e.g.*, β-galactosidase for generation of a dye, or luciferase for generation of luminescence) may be coupled to the sequence to identify position and/or quantity of a complementary sequence. Alternatively, where contemplated cDNA molecules are utilized for affinity isolation procedures, the cDNA may be coupled to a molecule that is known to have a high-affinity (*i.e.*, $K_d<10^{-4}mol^{-1}$) partner, such as biotin, or an oligo-histidyl tag. In another example, one or more phosphate groups may be exchanged for a radioactive phosphate group with a ^{32}P or ^{33}P isotope to assist in detection and quantification, where the radiolabeled cDNA is employed as a hybridization probe.

Therapeutic Methods Employing Prostate-Specific Or Testis-Specific Nucleic Acid Molecules, Polypeptides, and Antibodies

The invention includes methods of treating or preventing prostate-specific or testis-specific diseases. Therapies are designed to circumvent or overcome a prostate-specific or testis-specific gene defect, or inadequate or excessive prostate-specific or testis-specific gene expression, and thus modulate and possibly alleviate conditions involving defects in prostate-specific or testis-specific genes or proteins. In considering various therapies, it is understood that such therapies are, preferably, targeted to the affected or potentially affected organs, for example, the prostate or the testis. Reagents that are used to modulate prostate-specific or testis-specific biological activity can include, without limitation, full length prostate-specific or testis-specific polypeptides; prostate-specific or testis-specific *c*DNA, mRNA, or antisense RNA; prostate-specific or testis-specific antibodies; and any compound that modulates prostate-specific or

testis-specific polypeptide or nucleic acid molecule biological activity, expression, or stability.

Treatment or prevention of diseases resulting from a mutated prostate-specific or testis-specific gene is accomplished, for example, by replacing a mutant prostate-specific or testis-specific gene with a normal prostate-specific or testis-specific gene, administering a normal prostate-specific or testis-specific gene, modulating the function of a mutant prostate-specific or testis-specific protein, delivering normal prostate-specific or testis-specific protein to the appropriate cells, or altering the levels of normal or mutant prostate-specific or testis-specific protein. It is also possible to correct a prostate-specific or testis-specific gene defect to modify the physiological pathway (*e.g.*, an intracellular trafficking pathway) in which the prostate-specific or testis-specific protein participates.

To replace a mutant protein with normal protein, or to add protein to cells that do not express sufficient or normal prostate-specific or testis-specific protein, it may be necessary to obtain large amounts of pure prostate-specific or testis-specific protein from cultured cell systems in which the protein is expressed (see, *e.g.*, below). Delivery of the protein to the affected tissue can then be accomplished using appropriate packaging or administrating systems. Alternatively, small molecule analogs that act as prostate-specific or testis-specific molecule agonists or antagonists can be administered to produce a desired physiological effect (see below).

Gene therapy is another therapeutic approach for preventing or ameliorating diseases caused by prostate-specific or testis-specific gene defects. Nucleic acid molecules encoding wild type prostate-specific or testis-specific proteins can be delivered to cells that lack sufficient, normal prostate-specific or testis-specific biological activity (*e.g.*, cells carrying mutations in prostate-specific or testis-specific genes). The nucleic acid

molecules must be delivered to those cells in a form in which they can be taken up by the cells and so that sufficient levels of protein, to provide effective prostate-specific or testis-specific function, can be produced. Alternatively, for some prostate-specific or testis-specific mutations, it may be possible slow the progression of the resulting disease or to modulate prostate-specific or testis-specific activity by introducing another copy of a homologous gene bearing a second mutation in that gene, to alter the mutation, or to use another gene to block any negative effect.

Transducing retroviral, adenoviral, and adeno-associated viral vectors can be used for somatic cell gene therapy, especially because of their high efficiency of infection and stable integration and expression (see, *e.g.*, Cayouette *et al.*, *Human Gene Therapy* 8:423-430, 1997; Kido *et al.*, *Current Eye Research* 15:833-844, 1996; Bloomer *et al.*, *Journal of Virology* 71:6641-6649, 1997; Naldini *et al.*, *Science* 272:263-267, 1996; and Miyoshi *et al.*, *Proc. Natl. Acad. Sci., USA* 94:10319-1032, 1997). For example, the full length prostate-specific or testis-specific gene, or a portion thereof, can be cloned into a retroviral vector and expression can be driven from its endogenous promoter, from the retroviral long terminal repeat, or from a promoter specific for a target cell type of interest (such as aortic or other vascular cells). Other viral vectors that can be used include, for example, vaccinia virus, bovine papilloma virus, or a herpes virus, such as Epstein-Barr Virus (also see, for example, the vectors of Miller, *Human Gene Therapy* 15-14, 1990; Friedman, *Science* 244:1275-1281, 1989; Eglitis *et al.*, *BioTechniques* 6:608-614, 1988; Tolstoshev *et al.*, *Current Opinion in Biotechnology* 1:55-61, 1990; Sharp, *The Lancet* 337:1277-1278, 1991; Cornetta *et al.*, *Nucleic Acid Research and Molecular Biology* 36:311-322, 1987; Anderson, *Science* 226:401-409, 1984; Moen, *Blood Cells* 17:407-416, 1991; or Miller *et al.*, *Biotechnology* 7:980-990, 1989). Retroviral vectors are particularly well developed and have been used in

clinical settings (Rosenberg *et al.*, *N. Engl. J. Med* 323:370, 1990; Anderson *et al.*, U.S. Patent No. 5,399,346).

Gene transfer can also be achieved using non-viral means involving transfection *in vitro*, by means of any standard technique, including but not limited to, calcium phosphate, DEAE dextran, electroporation, protoplast fusion, and liposomes. Transplantation of normal genes into the affected tissues of a patient can also be accomplished by transferring a normal prostate-specific or testis-specific gene into a cultivatable cell type *ex vivo*, after which the cell (or its descendants) is injected into a targeted tissue. Another strategy for inhibiting prostate-specific or testis-specific function using gene therapy involves intracellular expression of an anti-prostate-specific or testis-specific antibody or a portion of an prostate-specific or testis-specific antibody. For example, the gene (or gene fragment) encoding a monoclonal antibody that specifically binds to prostate-specific or testis-specific polypeptide and inhibits its biological activity is placed under the transcriptional control of a tissue-specific gene regulatory sequence. Another therapeutic approach involves administration of recombinant prostate-specific or testis-specific polypeptide, either directly to the site of a potential or actual disease-affected tissue (for example, by injection) or systemically (for example, by any conventional recombinant protein administration technique). The dosage of a prostate-specific or testis-specific polypeptide depends on a number of factors, including the size and health of the individual patient but, generally, between about 0.006 mg/kg to about 0.6 mg/kg, inclusive, is administered per day to an adult in any pharmaceutically acceptable formulation.

Non-viral approaches can also be employed for the introduction of therapeutic DNA into cells predicted to be subject to diseases involving a prostate-specific or testis-specific disorder. For example, a prostate-specific or testis-specific nucleic acid molecule or an antisense nucleic acid

molecule can be introduced into a cell by lipofection (Felgner *et al.*, *Proc. Natl. Acad. Sci. USA* 84:7413, 1987; Ono *et al.*, *Neuroscience Letters* 17:259, 1990; Brigham *et al.*, *Am. J. Med. Sci.* 298:278, 1989; Staubinger *et al.*, *Methods in Enzymology* 101:512, 1983), asialoorosomucoid-polylysine conjugation (Wu *et al.*, *Journal of Biological Chemistry* 263:14621, 1988; Wu *et al.*, *Journal of Biological Chemistry* 264:16985, 1989), or, less preferably, micro-injection under surgical conditions (Wolff *et al.*, *Science* 247:1465, 1990).

Prostate-specific or testis-specific cDNA expression for use in gene therapy methods can be directed from any suitable promoter (*e.g.*, the human cytomegalovirus (CMV), simian virus 40 (SV40), or metallothionein promoters), and regulated by any appropriate mammalian regulatory element. For example, if desired, enhancers known to preferentially direct gene expression in specific cell types can be used to direct prostate-specific or testis-specific expression. The enhancers used can include, without limitation, those that are characterized as tissue- or cell-specific enhancers. Alternatively, if a prostate-specific or testis-specific genomic clone is used as a therapeutic construct (such clones can be identified by hybridization with prostate-specific or testis-specific cDNA, described above), regulation can be mediated by the cognate regulatory sequences, or, if desired, by regulatory sequences derived from a heterologous source, including any of the promoters or regulatory elements described above.

Antisense-based strategies can be employed to explore prostate-specific or testis-specific gene function and as a basis for therapeutic drug design. These strategies are based on the principle that sequence-specific suppression of gene expression (via transcription or translation) can be achieved by intracellular hybridization between genomic DNA or mRNA and a complementary antisense species. The formation of a hybrid RNA

duplex interferes with transcription of the target prostate-specific or testis-specific-encoding genomic DNA molecule, or processing, transport, translation, or stability of the target prostate-specific or testis-specific mRNA molecule.

Antisense strategies can be delivered by a variety of approaches. For example, antisense oligonucleotides or antisense RNA can be directly administered (*e.g.*, by intravenous injection) to a subject in a form that allows uptake into cells. Alternatively, viral or plasmid vectors that encode antisense RNA (or antisense RNA fragments) can be introduced into a cell *in vivo* or *ex vivo*. Antisense effects can be induced by control (sense) sequences; however, the extent of phenotypic changes are highly variable. Phenotypic effects induced by antisense effects are based on changes in criteria such as protein levels, protein activity measurement, and target mRNA levels.

For example, prostate-specific or testis-specific gene therapy can also be accomplished by direct administration of antisense prostate-specific or testis-specific mRNA to a cell that is expected to be adversely affected by the expression of wild-type or mutant prostate-specific or testis-specific polypeptides. The antisense prostate-specific or testis-specific mRNA can be produced and isolated by any standard technique, but is most readily produced by *in vitro* transcription using an antisense prostate-specific or testis-specific cDNA under the control of a high efficiency promoter (*e.g.*, the T7 promoter). Administration of antisense prostate-specific or testis-specific mRNA to cells can be carried out by any of the methods for direct nucleic acid molecule administration described above.

An alternative strategy for inhibiting prostate-specific or testis-specific function using gene therapy involves intracellular expression of an anti-prostate-specific or testis-specific antibody or a portion of an anti-

specific mRNA expression construct is undertaken to reverse or prevent the gene defect prior to the development of the full course of the disease.

The therapeutic methods of the invention are, in some cases, targeted to prenatal treatment. For example, a fetus found to have a prostate-specific or testis-specific mutation is administered a gene therapy vector including a normal prostate-specific or testis-specific gene, or administered a normal prostate-specific or testis-specific protein. Such treatment may be required only for a short period of time, or may, in some form, be required throughout such a patient's lifetime. Any continued need for treatment, however, is determined using, for example, the diagnostic methods described above. Also as discussed above, prostate-specific or testis-specific polypeptide or nucleic acid molecule abnormalities may be associated with diseases in adults, and thus, adults are subject to the therapeutic methods of the invention as well.

Additionally, prostate-specific or testis-specific polypeptides may be used to stimulate an immune system to assist in generating immunity against, for example, prostate cancer cells.

The methods of the present invention can be used to diagnose or treat the disorders described herein in any mammal, for example, humans, domestic pets, or livestock. Where a non-human mammal is treated or diagnosed, the prostate-specific or testis-specific polypeptide, nucleic acid molecule, or antibody employed is preferably specific for that species.

Identification of Molecules that Modulate Prostate-Specific Or Testis-Specific Polypeptide or Nucleic Acid Molecule Biological Activity or Whose Biological Activity is Modulated by Prostate-Specific Or Testis-Specific Polypeptides or Nucleic Acid Molecules

Isolation of prostate-specific or testis-specific cDNAs (as described herein) also facilitates the identification of molecules that increase or decrease prostate-specific or testis-specific polypeptide or nucleic acid molecule biological activity. Similarly, molecules whose activity is modulated by prostate-specific or testis-specific polypeptide or nucleic acid molecule biological activity can be identified. According to one approach, candidate molecules are added at varying concentrations to the culture medium of cells expressing prostate-specific or testis-specific mRNA. Prostate-specific or testis-specific biological activity is then measured using standard techniques. The measurement of biological activity can include, without limitation, the measurement of prostate-specific or testis-specific protein and nucleic acid molecule expression levels, response to androgens, or intracellular localization and trafficking.

If desired, the effect of candidate modulators on expression can also be measured at the level of prostate-specific or testis-specific protein production using the same general approach and standard immunological detection techniques, such as western blotting or immunoprecipitation with a prostate-specific or testis-specific-specific antibody (see below).

A test compound that is screened in the methods described above can be a chemical, be it naturally-occurring or artificially-derived. Such compounds can include, for example, polypeptides, synthesized organic molecules, naturally occurring organic molecules, nucleic acid molecules, and components thereof. Candidate prostate-specific or testis-specific modulators include peptide as well as non-peptide molecules (*e.g.*, peptide

or non-peptide molecules found, *e.g.*, in a cell extract, mammalian serum, or growth medium in which mammalian cells have been cultured).

Administration of Prostate-Specific Or Testis-Specific Polypeptides, Prostate-Specific Or Testis-Specific Nucleic Acid Molecules, and Modulators of Prostate-Specific Or Testis-Specific Polypeptide or Nucleic Acid Molecule Synthesis or Function

A prostate-specific or testis-specific protein, nucleic acid molecule, or modulator is administered within a pharmaceutically-acceptable diluent, carrier, or excipient, in unit dosage form to patients or experimental animals. Also, conventional pharmaceutical practice is employed to provide suitable formulations or compositions in which to administer neutralizing prostate-specific or testis-specific antibodies or prostate-specific or testis-specific-inhibiting compounds (*e.g.*, a prostate-specific or testis-specific antisense molecule or a prostate-specific or testis-specific dominant negative mutant) to patients suffering from a prostate-specific or testis-specific disease, such as prostate cancer, testicular cancer, benign hyperplasia of the prostate, or developmental defects of the prostate or testis. Administration can begin before or after the patient is symptomatic.

Any appropriate route of administration can be employed, for example, administration can be parenteral, intravenous, intra-arterial, subcutaneous, intramuscular, intracranial, intraorbital, ophthalmic, intraventricular, intracapsular, intraspinal, intracisternal, intraperitoneal, intranasal, inhalation to deep lung, aerosol, by suppositories, oral, or topical (*e.g.* by applying an adhesive patch carrying a formulation capable of crossing the dermis and entering the bloodstream). Preferably, the administration is local to the afflicted tissue, such as prostate or testis tissue. Therapeutic formulations can be in the form of liquid solutions or suspensions; for oral administration, formulations can be in the form of

tablets or capsules; and for intranasal formulations, in the form of powders, nasal drops, or aerosols. Any of the above formulations may be a sustained-release formulation.

Methods that are well known in the art for making formulations are found, for example, in *Remington's Pharmaceutical Sciences*, (18^{th} edition), ed. A. Gennaro, 1990, Mack Publishing Company, Easton, PA. Formulations for parenteral administration can, for example, contain excipients; sterile water; or saline; polyalkylene glycols, such as polyethylene glycol; oils of vegetable origin; or hydrogenated napthalenes. Sustained-release, biocompatible, biodegradable lactide polymer, lactide/glycolide copolymer, or polyoxyethylene-polyoxypropylene copolymers can be used to control the release of the compounds. Other potentially useful parenteral delivery systems for prostate-specific or testis-specific modulatory compounds include ethylene-vinyl acetate copolymer particles, osmotic pumps, implantable infusion systems, and liposomes. Formulations for inhalation can contain excipients, for example, lactose, or can be aqueous solutions containing, for example, polyoxyethylene-9-lauryl ether, glycocholate, and deoxycholate, or can be oily solutions for administration in the form of nasal drops, or as a gel.

Prostate-Specific Or Testis-Specific Fragments

Polypeptide fragments that include various portions of prostate-specific or testis-specific proteins are useful in identifying the domains important for their biological activities, such as protein-protein interactions and transcription. Methods for generating such fragments are well known in the art (see, for example, Ausubel *et al.*, *supra*), using the nucleotide sequences provided herein. For example, a prostate-specific or testis-specific protein fragment can be generated by PCR amplifying a desired prostate-specific or testis-specific nucleic acid molecule fragment using

oligonucleotide primers designed based upon the prostate-specific or testis-specific nucleic acid sequences. Preferably, the oligonucleotide primers include unique restriction enzyme sites that facilitate insertion of the amplified fragment into the cloning site of an expression vector (*e.g.*, a mammalian expression vecoor, see above). This vector can then be introduced into a cell (*e.g.*, a mammalian cell; see above) by artifice, using any of the various techniques known in the art such as those described herein, resulting in the production of a prostate-specific or testis-specific polypeptide fragment in the cell containing the expression vector.

Prostate-specific or testis-specific polypeptide fragments (*e.g.*, chimeric fusion proteins) can also be used to raise antibodies specific for various regions of prostate-specific or testis-specific polypeptides. Preferred prostate-specific or testis-specific fragments include, without limitation, fragments including the N-terminal domain of STMP1 (amino acids 1-200), the P5CR domain, and fragments thereof.

Synthesis of Prostate-Specific Or Testis-Specific Proteins, Polypeptides, and Polypeptide Fragments

Those skilled in the art of molecular biology will understand that a wide variety of expression systems can be used to produce recombinant prostate-specific or testis-specific proteins. The precise host cell used is not critical to the invention. The prostate-specific or testis-specific proteins can be produced in a prokaryotic host (*e.g.*, *E. coli*) or in a eukaryotic host (*e.g.*, *S. cerevisiae*, insect cells such as Sf9 cells, or mammalian cells such as COS, NIH 3T3, CHO, or HeLa cells). These cells are commercially available from, for example, the American Type Culture Collection, Rockville, MD (see also Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY, 1998). The method of transformation and the choice of expression vehicle (*e.g.*, expression

vector) will depend on the host system selected. Transformation and transfection methods are described, *e.g.*, in Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY, 1998, and expression vehicles can be chosen from those provided, *e.g.* in Pouwels *et al.*, *Cloning Vectors: A Laboratory Manual*, 1985, Supp. 1987).

The characteristics of prostate-specific or testis-specific nucleic acid molecules are analyzed by introducing such genes into various cell types or using *in vitro* extracellular systems. The function of prostate-specific or testis-specific proteins produced in such cells or systems are then examined under different physiological conditions. Also, cell lines can be produced that over-express the prostate-specific or testis-specific gene product, allowing purification of prostate-specific or testis-specific proteins for biochemical characterization, large-scale production, antibody production, and patient therapy.

The polypeptides of the invention may be produced *in vivo* or *in vitro*, and may be chemically and/or enzymatically modified. The polypeptides can be isolated from prostate tissue or prostate cancer cells that may or may not be in a hormone dependent state. Alternatively, and especially where larger amounts (*i.e.*, >10mg) are desirable, recombinant production (*e.g.*, in a bacterial, yeast, insect cell, or mammalian cell system) may advantageously be employed to generate significant quantities of prostate-specific or testis-specific polypeptides.

Furthermore, recombinant production not only offers a more economical strategy to produce the polypeptides of the invention, but also allows specific modification in the amino acid sequence and composition to tailor particular biochemical, catalytic and physical properties. For example, where increased solubility of is desirable, one or more hydrophobic amino acids may be replaced with hydrophilic amino acids.

Alternatively, where reduced or increased catalytic activity is required, one or more amino acids may be replaced or eliminated.

In still another example, the polypeptides of the invention can be synthesized as fusion proteins including, for example, fusions with enzymatically active partners (*e.g.*, for dye formation or substrate conversion) and fluorescent partners such as GFP, EGFB, BFP, etc.

With respect to chemical and enzymatic modifications of contemplated polypeptides, it is many modifications are appropriate, including addition of mono-, and bifunctional linkers, coupling with protein- and non-protein macromolecules, and glycosylation. For example, mono- and bifunctional linkers are especially advantageous where polypeptides are immobilized to a solid support, or covalently coupled to a molecule that enhances immunogenicity of contemplated polypeptides (*e.g.*, KLH, or BSA conjugation). Alternatively, the polypeptides may be coupled to antibodies or antibody fragments to allow rapid retrieval of the polypeptide from a mixture of molecules. Further couplings include covalent and non-covalent coupling of polypeptides with molecules that prolong the serum half-life and/or reduce immunogenicity such as cyclodextranes and polyethylene glycols.

Use of Prostate-Specific Or Testis-Specific Antibodies

Antibodies to prostate-specific or testis-specific proteins are used to detect prostate-specific or testis-specific proteins or to inhibit the biological activities of prostate-specific or testis-specific proteins. For example, a nucleic acid molecule encoding an antibody or portion of an antibody can be expressed within a cell to inhibit prostate-specific or testis-specific function. In addition, the antibodies can be coupled to compounds, such as radionuclides and liposomes for diagnostic or therapeutic uses. Antibodies that inhibit the activity of a prostate-specific or testis-specific polypeptide

can also be useful in preventing or slowing the development of a disease caused by inappropriate expression of a wild type or mutant prostate-specific or testis-specific gene. For example, the antibodies of the invention may be utilized to localize and locally quantify disease-specific markers in prostate or testis tissue sections, e.g, in prostate or testicular cancer.

Detection Of Prostate-Specific Or Testis-Specific Gene Expression

As noted, the antibodies described above can be used to monitor prostate-specific or testis-specific protein expression. *In situ* hybridization of RNA can be used to detect the expression of prostate-specific or testis-specific genes. RNA *in situ* hybridization techniques rely upon the hybridization of a specifically labeled nucleic acid probe to the cellular RNA in individual cells or tissues. Therefore, RNA *in situ* hybridization is a powerful approach for studying tissue- and temporal-specific gene expression. In this method, oligonucleotides, cloned DNA fragments, or antisense RNA transcripts of cloned DNA fragments corresponding to unique portions of prostate-specific or testis-specific genes are used to detect specific mRNA species, *e.g.*, in the tissues of animals, such as mice, at various developmental stages, or to monitor tumor progression. Other gene expression detection techniques are known to those of skill in the art and can be employed for detection of prostate-specific or testis-specific gene expression.

Identification of Additional Prostate-Specific Or Testis-Specific Genes

Standard techniques, such as the polymerase chain reaction (PCR) and DNA hybridization, as well as the SSH and other techniques described herein, can be used to clone prostate-specific or testis-specific homologues in other species and other prostate-specific or testis-specific genes in

humans. Prostate-specific or testis-specific genes and homologues can be readily identified using low-stringency DNA hybridization or low-stringency PCR with human prostate-specific or testis-specific probes or primers. Degenerate primers encoding human prostate-specific or testis-specific or human prostate-specific or testis-specific amino acid sequences can be used to clone additional prostate-specific or testis-specific genes and homologues by RT-PCR.

Additional prostate-specific or testis-specific genes include genes expressed during various growth and developmental phases of the diseased prostate or testis, e.g., those involved in prostate cancer, benign prostatic hyperplasia, or testicular cancer, and genes expressed as a result of a drug regimen.

Construction of Transgenic Animals and Knockout Animals

Characterization of prostate-specific or testis-specific genes provides information that allows prostate-specific or testis-specific knockout animal models to be developed by homologous recombination. Preferably, a prostate-specific or testis-specific knockout animal is a mammal, most preferably a mouse. Similarly, animal models of prostate-specific or testis-specific overproduction can be generated by integrating one or more prostate-specific or testis-specific sequences into the genome of an animal, according to standard transgenic techniques. Moreover, the effect of prostate-specific or testis-specific gene mutations (*e.g.*, dominant gene mutations) can be studied using transgenic mice carrying mutated prostate-specific or testis-specific transgenes or by introducing such mutations into the endogenous prostate-specific or testis-specific gene, using standard homologous recombination techniques.

A replacement-type targeting vector, which can be used to create a knockout model, can be constructed using an isogenic genomic clone, for

example, from a mouse strain such as 129/Sv (Stratagene Inc., LaJolla, CA). The targeting vector can be introduced into a suitably-derived line of embryonic stem (ES) cells by electroporation to generate ES cell lines that carry a profoundly truncated form of a prostate-specific or testis-specific gene. To generate chimeric founder mice, the targeted cell lines are injected into a mouse blastula-stage embryo. Heterozygous offspring can be interbred to homozygosity. Prostate-specific or testis-specific knockout mice provide a tool for studying the role of prostate-specific or testis-specific polypeptides and nucleic acid molecules in embryonic development and in disease. Moreover, such mice provide the means, *in vivo*, for testing therapeutic compounds for amelioration of diseases or conditions involving a prostate-specific or testis-specific polypeptide or nucleic acid molecule-dependent or prostate-specific or testis-specific polypeptide or nucleic acid molecule-affected pathway.

Animal Models

The prostate-specific and testis-specific polypeptides, antisense compounds, etc., of the invention can also be used in conjunction with animal models of prostate or testis disorders, to test the therapeutic, diagnostic, and screening methods of the invention. An exemplary prostate cancer model in transgenic mice is called TRAMP, in which the SV40 large T antigen is targeted to the prostate (Greenberg et al., PNAS 92, 3439-3443, 1995). Another test system is the CWR22 (androgen-dependent) and CWR22R (androgen-independent) xenografts, as known in the art and as described herein. Growth, PSA secretion, metastasis, etc. of these xenografts could be monitored in the presence and absence of the prostate-specific or testis-specific polypeptides, nucleic acid molecules, and other compounds of the invention. Other animal models, for example, animal

models of other forms of cancer, or immunocompromised animals, e.g., nude mice, may also be used.

The following Examples will assist those skilled in the art to better understand the invention and its principles and advantages. It is intended that these Examples be illustrative of the invention and not limit the scope thereof.

EXAMPLE 1

Suppression Subtraction Of Prostate- And Testes-Specific Genes And Subcloning Into Pzero

cDNA derived from poly(A)+ RNA of 10 different normal human tissues were subtracted against normal human prostate cDNA using suppression subtraction hybridization (SSH) (Diatchenko, L. et al., *Proc. Natl. Acad. Sci.* USA 93, 6025-6030, 1996) and the resulting cDNA fragments were cloned into an appropriate vector. SSH was performed as described (Clontech PCR-Select Cloning Kit) using prostate poly(A)+ RNA against a pool of poly(A)+ RNA obtained from ten normal human tissues (heart, brain, placenta, lung, liver, skeletal muscle, kidney, spleen, thymus, and ovary). Upon secondary PCR amplification (12 cycles), the reactions were extracted with phenol/chloroform, the DNA with ethanol, and the pellets washed once with 70% ethanol. After drying, the DNA pellet was dissolved in 0.2XTE or MQ dH_2O and cut with RsaI in a 20 µl reaction for 2 hrs at 37°C to excise adaptors. After digestion, the reactions were run on a 1.5% agarose gel, with molecular size markers on one side, at 5 V/cm, 40 min. Care was taken not to expose the gel to short wavelength UV light. The adapter bands were excised, and the gel was run at 5 V/cm for 15 min in a reversed electric field to concentrate the cDNA bands.

The gel was visualized (long wave UV light) and the amplified cDNAs, ranging in size between 100 bp-1kB, were excised. The DNA was purified using the QAIEX gel DNA purification kit. The purified DNA was cloned into EcoRV-cut, dephosphorylated pZERO (Invitrogen). Ligation reactions were performed in 10 µl final volume in the presence of 5% PEG, 1X T4 Ligase buffer at 37°C overnight and a 1/5 dilution of 1µl of the ligation mix (PSL) was transformed into DH10B electrocompetent cells (>10^{10} efficiency) or equivalent. Colonies were picked and the presence of cDNA inserts was confirmed. To that end, PCR was performed with T7 and SP6 primers directly from the colonies. 10% of the reactions were run on a 1.5% agarose gel to visualize amplified products. The colonies with inserts were grown and glycerol stocks (15%) were prepared and stored at –80°C.

EXAMPLE 2

Reverse Northern Blot And Sequence Analyses

To clone androgen-responsive genes represented in the PSL, the reverse northern technique was used (Hedrick, S.M. et al., *Nature* 308, 149-153, 1984; Sakaguchi, N. et al., *EMBO J* 5: 2139-2147, 1986). In this procedure, RNA made from two populations of cells that are to be compared is used to make cDNA probes that are then hybridized to two identical arrays of clones. To that end, PSL clones were amplified by PCR and spotted on nylon filters in 96-well format to generate two identical blots for each set of 92 clones (the remaining four spots were used for positive and negative controls). To make the probes, the androgen-responsive prostate cancer cell line LNCaP was used (Horoszewicz, J.S. et al., *Cancer Res.* 43, 1809-1818, 1983) and was either left untreated (the (-) probe) or treated with the synthetic androgen R1881 for 24 hours (the (+) probe). Poly(A)+ RNA was isolated from these cells and was used to make

the ^{32}P-labeled probes. After hybridization with the (-) and (+) probes, clones that showed differential hybridization were selected for further analysis, i.e., confirmation by a secondary reverse northern blot, and northern blotting.

Reverse northern screening on the cDNA clones was done essentially as described previously (Hedrick, S.M. et al., *supra*; Sakaguchi, N. et al., *supra*) with some modifications. DNA (approximately 400 ng) from PCR amplification in step 6 was diluted in 200 µl of 0.4M NaOH, 10 mM EDTA and mixed well by pipetting. After incubation at 95°C for 5-10 minutes, the tubes were chilled on ice. Denatured DNA was blotted on two separate pieces of Zeta Probe GT+ membrane (Bio-Rad) using a dot-blot apparatus (Bio-Rad). Positive (Prostate specific antigen (PSA) cDNA) and negative (glyceraldehyde 3-phosphate dehydrogenase (G3PDH) cDNA) controls were included on each blot (bottom right) in duplicate. The membranes were rinsed with 2XSSC, air dried, and then baked at 80°C for 30 minutes. An exemplary reverse northern analysis is shown in Figure 1. Note that there was a substantial increase in PSA hybridization in the (+) blot (probe prepared from cells that have been stimulated by androgens) compared with the (-) blot (probe prepared from unstimulated cells), whereas there was no significant change in hybridization of G3PDH between the two blots. Arrowheads indicate the positive clones identified in this experiment.

To verify the tissue-specific nature of the isolated sequences, positive clones were tested in a standard northern blot against RNA preparations of multiple non-prostate tissue samples. Figure 2 shows a multiple tissue northern blot using NKX3A as a probe, to show an exemplary tissue expression pattern seen in the positive clones. Lanes 1-10, and 12-16 are RNA preparations from non-prostate tissues, lane 11 is a RNA preparation from prostate, lane 12 is a RNA preparation from testis.

Twelve clones with no significant homology to known sequences (by BLAST analysis) were isolated from prostate tissue and LNCaP cells. SEQ ID NOs: 1-9 were identified as androgen-responsive differentially-expressed genes in the prostate, while SEQ ID NOs: 10-12 were identified as androgen-responsive differentially-expressed genes in LNCaP cells.

EXAMPLE 3

Isolation And Characterization Of The STMP1 Gene And mRNA

A normal prostate cDNA library was screened by 5'- and 3'-RACE analysis, and resulted in the full-length cDNA for L74. Since computer-aided secondary structure prediction of the deduced amino acid sequence of L74 suggested the presence of a six-transmembrane domain in its C-terminal half, L74 was renamed Six-Transmembrane Protein of Prostate 1 (STMP1).

When the full-length *STMP1* cDNA was used in BLAST analysis, it was found to match a BAC clone (GenBank accession # AC002064) except for a 313 bp repetitive unit in the 3' UTR region, thereby identifying it as the *STMP1* gene and localizing it to Chr7q21. The repetitive region is likely to be a cloning or sequencing artifact of the BAC clone. Computational exon/intron junction analysis and alignment of the full-length cDNA sequence with the BAC clone revealed that *STMP1* gene is composed of six exons and five introns (Figure 4A). The transcription start site, the location and size of the exons and introns, and the location of the partial cDNA clone L74 (black box) are indicated. The start (atg) and stop codons (tga), as well as the putative polyadenylation signal (pA) are also indicated. The first two exons are short, non-coding exons of 83 and 61 bp, whereas exons 3-6 encode the open reading frame (ORF) and are 525, 528, 165, and 3281 bp long, respectively (Figure 4C). The *STMP1* gene spans around 26 kb, which is in part due to the extremely large size of intron 2 (12713 bp).

There are three different predicted promoters within 4 kb upstream of the *STMP1* initiation codon, none of which has any significant TATA or CAAT box consensus sequences, suggesting that *STMP1* is transcribed from a TATA-less promoter.

The *STMP1* cDNA (GenBank accession # AY008445) has a predicted 5′ untranslated region (5′UTR) of approximately 1 kb (deduced by RACE analysis) and an unusually long 3′UTR of approximately 4 kb that comprises ~77% of the total cDNA sequence. The ORF starts within the 3rd exon and is predicted to encode a 490 amino-acid protein (Figure 4B). A search for protein motifs identified six predicted transmembrane domains in the C-terminal half of STMP1 starting at F209 (Figures 4B and 4E). Only the cDNA sequence surrounding the ORF is indicated. The exon-intron junctions are indicated and the location of the predicted transmembrane domains are highlighted (TM 1-6) (Figure 4B). The stop codon is indicated with an asterisk. STMP1 has two alternatively spliced forms, shown in Figures 4F-4K, which lead to two predicted isoforms of the protein.

EXAMPLE 4

STMP1 Belongs To A New Subfamily Of Six-Transmembrane Domain Proteins

BLAST analysis of GenBank with the predicted STMP1 amino acid sequence identified two independent ESTs and STEAP, a recently discovered cell membrane protein enriched in prostate for expression. An alignment of these sequences, obtained by Clustal and GenDoc programs, is shown in Figure 5. Completely conserved residues are shaded in black; residues that are conserved in two or three of the sequences are shaded light and dark gray, respectively. This alignment suggested that while the EST

BAA91839 cDNA may be close to full-length, BAB15559 cDNA may represent a partial sequence.

The sequences of two proteins related to STMP1 were determined (Figures 4L and 4M, STMP2 and STMP3, respectively). The STMP2 and STMP3 sequences contain the EST sequences. The GFP-fusion of STMP2 gives similar localization as STMP1. Both STMP2 and STMP3 are more widely distributed and have higher levels in some tissues other than the prostate. For example, STMP2 has the highest expression in the placenta and the lung, and is also highly expressed in the heart, liver, prostate, and testis, while STMP3 has the highest expression in the liver, and is also highly expressed in the heart, placenta, lung, kidney, pancreas, prostate, testis, small intestine, and colon.

The sequence similarity between STMP1 and STEAP is limited and not significant before residue 210 of STMP1 where the predicted six-transmembrane coding domain starts. This suggests that the N-terminal region is structurally and functionally related among STMP proteins, forming a six-transmembrane protein subfamily that is distinct from STEAP.

EXAMPLE 5

STMP1 Expression Is Highly Enriched In Prostate

The expression profile of *STMP1* was then determined in various human tissues by Northern analysis, in which a multiple tissue Northern blot was hybridized to the *STMP1* probe (see Materials and Methods). As shown in Figure 6A, *STMP1* hybridized to a major mRNA species of 6.5 kb, and three minor mRNA species of 2.2, 4.0, and 4.5 kb in the prostate tissue. The stronger hybridization that is observed with G3PDH in the heart and skeletal muscle samples is due to its higher expression in these tissues. The lanes represent: 1.Heart, 2. Brain, 3. Placenta, 4. Lung, 5.

Liver, 6. Skeletal Muscle, 7. Kidney, 8. Pancreas, 9. Spleen, 10. Thymus, 11. Prostate, 12. Testis, 13. Ovary, 14. Small Intestine, 15. Colon, 16. Peripheral blood leukocyte. The location of the full-length 6.5 kb mRNA, as well as the lower molecular weight *STMP1* species are indicated by arrows to the left of the figure. There was 15-20-fold lower mRNA expression of the 6.5 kb band in the heart, brain, kidney, pancreas, and ovary, compared to prostate, and no detectable expression in other tissues. In contrast, the three lower molecular weight species, encoded by alternatively spliced forms of *STMP1*, were only detectable in the prostate. Hybridization with a glyceraldehyde 3-phosphate dehydrogenase (G3PDH) cDNA probe resulted in approximately similar signals in all lanes, except for the heart and skeletal muscle where G3PDH is known to be more abundant compared with other tissues. These data show that *STMP1* expression is high in the prostate, although expression can be seen in other tissues, and that *STMP1* has isoforms that are restricted to the prostate for expression.

EXAMPLE 6

Characterization Of STMP1 Expression

Since androgen is a major hormonal stimulus for the normal prostate gland and for early stage prostate cancer, the possible androgen regulation of *STMP1* was assessed by Northern analysis in the androgen-responsive prostate cancer cell line LNCaP. Cells were either left untreated or treated with the synthetic androgen R1881 (10^{-8} M) with increasing amounts of time (hours) as indicated (Figure 6B), harvested, and total RNA isolated and used in Northern analysis with *STMP1* cDNA as probe. The same membrane was also probed for the androgen-dependent gene PSA. Relative induction of mRNA accumulation is indicated at the bottom of the lanes, as determined by phosphorimager analysis (Molecular Dynamics).

The CWR22 xenograft was grown in nude mice and tumor samples were collected either before (t=0) or 1, 2, or 4 weeks after castration. Total RNA was isolated and was then used in Northern analysis with the same probes. Ethidium bromide-stained 18S RNA is shown as a control for RNA integrity and loading. At 6 h, there was an approximately 25% increase in *STMP1* expression, which was lost by 24 h, with a final 20% decrease observed at 48 h compared with basal levels. In contrast, the mRNA accumulation of the androgen-regulated gene PSA dramatically increased upon androgen stimulation in a time-dependent manner, as expected, reaching approximately 22-fold higher levels by 48 hours. Relative induction of *STMP1* mRNA accumulation is indicated at the bottom of the lanes determined by phosphorimager analysis. As is shown in Figure 6B, STMP1 displayed similar expression levels in untreated and R1881-treated LNCaP cells, indicating that *STMP1* expression is not significantly regulated by androgens in LNCaP cells.

To determine the possible androgenic regulation of *STMP1* expression in an *in vivo* setting, the androgen-dependent xenograft model CWR22, which is derived from a primary human prostate tumor was used (Wainstein, M. A. et al., *Cancer Res* 54, 6049-6052, 1994). Since they are androgen-dependent for growth, the CWR22 tumors in nude mice display marked regression upon castration and may regress completely. CWR22 xenografts were grown in nude mice in the presence of a sustained release testosterone pellet. After the tumors had grown, the mice were castrated, the testosterone pellets were removed, and the regressing tumors were collected at 1, 2, or 4 weeks post-castration. Total RNA was prepared from these tumor samples and used in Northern analysis. As shown in Figure 6B, similar to the obsevations in LNCaP cells, *STMP1* mRNA accumulation in the CWR22 tumors showed no significant change upon castration and was not affected by the presence of androgens (note that

there is underloading of RNA for CWR22 2wk sample). In contrast, the mRNA accumulation of the androgen-regulated gene PSA was dramatically decreased upon castration, dropping to approximately 16% of pre-castrate levels by two weeks post-castration. These results are consistent with the findings in LNCaP cells and suggest that *STMP1* expression is not significantly regulated by androgens in prostate cancer cells. *STMP1* expression was substantially lower in the CWR22 tumors compared with LNCaP cells.

The expression profile of *STMP1* was also analyzed in the androgen-independent prostate cancer cell lines PC3 and DU145, as well as in four independent, relapsed derivatives of CWR22 tumors, named CWR22R (Nagabhushan, M. et al., *Cancer Res* 56, 3042-3046, 1996), representative of advanced prostate cancer (Figure 6C). LNCaP (in the presence (+) or absence (-) of R1881 (10^{-8} M)), PC-3, or DU-145 cells were grown and total RNA was isolated. Four independent lines of the androgen independent human prostate cancer xenograft CWR22R, were grown in nude mice, tumors were collected, and total RNA was isolated and used in Northern analysis with *STMP1* or the androgen target gene *NKX3.1* cDNAs as probes. Ethidium bromide-stained 18S RNA is shown as a control for RNA integrity and loading. The relative induction of *STMP1* and *NKX3.1* mRNA accumulation is indicated at the bottom of the lanes determined by phosphorimager analysis (Molecular Dynamics). As is shown in Figure 6C, *STMP1* expression was high in LNCaP cells and did not significantly change in response to R1881 treatment compared with a ~9-fold induction of the androgen target gene *NKX3.1*. There was no *STMP1* expression in the androgen-independent prostate cancer cell lines PC-3 or DU-145, as was the case for *NKX3.1*. In contrast, there was significant *STMP1* expression in tumors from all four independent CWR22R xenograft lines tested, ranging between ~30-60% of that observed in LNCaP cells. A

similar overexpression pattern was also observed for *NKX3.1* (Figure 6C) consistent with previous findings (Korkmaz, K. S. et al., *Gene* 260, 25-36, 2000).

An interesting property of *STMP1* expression profile is that even though it is expressed at low levels in the androgen dependent CWR22 xenograft, it is highly expressed in the relapsed CWR22R which is androgen receptor (AR) positive, but is not responsive to androgens. This indicates that STMP1 expression is deregulated once the prostate tumor progresses from an androgen-dependent to an androgen-independent phase. In addition, *STMP1* is not expressed in the AR-negative prostate cancer cell lines PC-3 and DU-145, but is expressed at high levels in the AR-positive cell line LNCaP and the CWR22 and CWR22R xenografts. Thus, expression of *STMP1* is correlated with the presence of a functional AR in the cell.

It has been known for over 50 years that androgens play a key role both in the development and maintenance of the normal prostate and the initiation and progression of prostate cancer. Androgen withdrawal results in involution of both the normal prostate gland as well as a prostate tumor in the early stages of the disease that is still androgen dependent. Consequently, androgen withdrawal is commonly used as treatment to reverse tumor growth. However, in the case of the prostate tumor, after a few months or years, the tumor recurs in almost all cases in an androgen-independent state. At this point there is no effective therapy and prognosis for survival is extremely poor. Since *STMP1* is overexpressed during this later androgen-insensitive state, it will be a useful tool in diagnostic and therapeutic applications for prostate cancer.

These data indicate that *STMP1* expression is deregulated once prostate cancer progresses from an androgen-dependent to an androgen-independent state.

EXAMPLE 7

Intracellular Localization Of STMP1

To gain insight into the intracellular localization pattern of *STMP1*, a green fluorescent protein (GFP)-STMP1 fusion protein was generated. The use of such GFP chimeric proteins has recently become a standard method to assess intracellular localization and dynamics of proteins. COS-1 cells were transiently transfected with GFP-STMP1, fixed and processed for confocal microscopy as described in Materials and Methods.

A series of 11 confocal sections along the z-axis were collected through a single cell at nominal 100 nm intervals. Three of the consecutive sections and the projection of all 11 sections are shown in Figure 7A. Arrows indicate tubular-vesicular structures (VTS) in different sizes, shapes, and locations (Bar=5μm). In all 11 z-plane sections, GFP-STMP1 showed bright juxtanuclear distribution pattern, characteristic of the Golgi complex. Additionally, GFP-STMP1 was dispersed in spots of variable size throughout the cytoplasm and at the cell periphery (z-7, projection). Some of these bright fluorescent spots were tubular (z-6, arrow and Figure 8) or vesicular (z-5, arrow) in morphology.

To determine more directly whether GFP-STMP1 was localized to the Golgi complex, we compared its intracellular distribution with those of two well characterized Golgi markers, the medial Golgi enzyme mannosidase II (ManII) (Rabouille, C. et al., *J Cell Sci* 108, 1617-1627, 1995) and the coat protein β-COP (Pepperkok, R. et al., *Cell* 74, 71-82, 1993). COS-1 cells were transfected with GFP-STMP1, fixed, labeled with the appropriate primary and secondary antibodies and imaged by confocal laser scanning microscopy. Green GFP-STMP1 fluorescence and red (Texas Red-labeled secondary antisera) β-COP and ManII fluorescence were detected by confocal laser microscopy. Panels to the right show the

overlay images with yellow/orange staining indicating the regions of colocalization. Bars=5μm. As shown in Figure 7B, the distribution of GFP-STMP1 extended throughout the Golgi complex, as evidenced by significant colocalization with both ManII and β-COP. However, some areas of non-overlap between the GFP-STMP1 and both Golgi markers were observed suggesting that STMP1, at least in part, is differentially localized within the Golgi complex compared with these two markers.

Since GFP-STMP1 was associated with VTS (Figure 7A and Figure 8), more specific localization of GFP-STMP1 to the trans-Golgi network (TGN), an important site for the sorting of proteins destined to the plasma membrane, secretory vesicles, or lysosomes (Farquhar, M. G. & Palade, G. E. *Trends Cell Biol* 8, 2-10, 1998; Mellman, I. & Warren, G., *Cell* 100, 99-112, 2000; Lemmon, S. K. & Traub, L. M., *Curr Opin Cell Biol* 12, 457-466, 2000) was assessed. An antibody against TGN46, a TGN resident protein that shuttles between the TGN and the plasma membrane (Prescott AR, et al., *Eur J Cell Biol* 72, 238-246, 1997; Ponnambalam, S. et al., *J Cell Sci.* 109, 675-685, 1996), was used in immunoflourescence microscopy experiments as above. As shown in Figure 7B, GFP-STMP1 extensively colocalized with TGN46, greater than that observed with ManII and β-COP, suggesting that in the Golgi complex, STMP1 is primarily localized to the TGN. Note that the images with TGN46 were obtained with lower objective power.

EXAMPLE 8

STMP1 Shuttles Between The Golgi And The Plasma Membrane And Colocalizes To The Early Endosomes

The dynamic properties and intracellular trafficking of GFP-STMP1 were studied using confocal time-lapse imaging in living cells. COS-1 cells were transiently transfected with GFP-STMP1 and, 16 h after transfection,

12 consecutive images were collected from live cells every 20s at 37°C by confocal laser scanning microscope (Figure 8). The upper panel shows a VTS extending out and retracting back to the Golgi body (white arrows). In the middle panel and the first image in the lower panel (160s), red arrows indicate the translocation of a VTS from the Golgi body to the cell periphery. In the lower panel, yellow arrows point to the movement of a VTS from the edge of the cell towards the Golgi body. Note that the results shown are representative of multiple time-lapse analyses and the changes in the images are not due to movement from the plain of focus. Bar=5μm.

As shown in Figure 8, some VTS were found to be detaching and some to be associating with the Golgi complex. The VTS were highly dynamic and pleiomorphic in size. Some of the VTS followed straight or curvilinear paths, some moved in a stop-and-go fashion, and some showed saltatory movements. The VTS indicated at the top panel (white arrows) extended away from and then retracted back to the Golgi. The VTS in the middle panel and the first image in the lower panel (red arrows) detached from the Golgi complex, paused, and then moved towards the cell periphery until it disappeared at the cell edge suggesting that STMP1 is associated with the secretory pathway. The VTS in the lower panel (yellow arrow) moved from the cell periphery towards the Golgi body suggesting that STMP1 is localized to the endocytic pathway.

EXAMPLE 9

Colocalization Of GFP-STMP1 With The Early Endosomal Marker EEA1

To probe whether GFP-STMP1 was associated with the endocytic pathway, the intracellular distribution of GFP-STMP1 was compared with that of the early endosome protein EEA1 (Stenmark, H. et al., *J Biol Chem* 271, 204048-204054, 1996). COS-1 cells were transfected with GFP-STMP1, fixed, immunostained with EEA1 antibodies and observed by

confocal laser scanning microscopy. Green GFP-STMP1 fluorescence and red (Texas Red-labeled secondary antiserum) EEA1 fluorescence were detected by confocal laser microscopy. The panel to the right shows the overlay images with yellow/orange staining indicating the regions of colocalization. Arrows indicate examples of the VTS in the cell periphery which contain both EEA1 and STMP1. Bar=5μm. As shown in Figure 9, EEA1 manifested a similar intracellular distribution in both transfected and untransfected cells. Furthermore, GFP-STMP1 significantly colocalized with EEA1 both in the cell periphery and also in the perinuclear area (Figure 9, arrows) suggesting that STMP1 is associated with early endosomes and the endocytic pathway.

EXAMPLE 10

Isolation And Characterization of the SSH9 Gene And mRNA

The *SSH9* gene was identified and mapped (Figure 10). The predicted promoter site, the transcription start site, and the location and size of the exons and introns are indicated. The start and stop codons, as well as two polyadenylation signals, leading to two alternatively spliced transcripts, are also indicated. Figures 11A-C show the nucleotide and predicted amino acid sequence of SSH9, as well as the predicted promoter sequence and exon-intron boundaries.

The expression profile of *SSH9*, determined in various human tissues by Northern analysis (Figure 12C), revealed that the 0.7 kb splice variant of *SSH9* was highly testis-specific, while the 1.4 kb transcript was expressed in both prostate and testis.

The androgen regulation of *SSH9* was examined in LNCaP cells and in CWR22 xenografts (Figure 12A) revealed that *SSH9* is not regulated in LNCaP cells, but is regulated in CWR22 xenografts. The expression

profile of *SSH9* was also examined in the androgen-independent prostate cancer cell lines PC3 and DU145, and in CWR22R cells (Figure 12B).

EXAMPLE 11

Isolation And Characterization of the *PSL22* Gene And mRNA

The *PSL22* gene was identified and mapped (Figure 13). The location and size of the exons and introns, the location of the partial cDNA clone (black box), as well as the alignment of the full-length cDNA clone with GenBank Accession Nos. AC008551 and AC011449, are indicated. Figures 14A-C show the nucleotide sequence of the ORF, cDNA and predicted amino acid sequence, as well as the predicted promoter, exon, and UTR sequences of PSL22.

BLAST analysis of GenBank with the predicted PSL22 amino acid sequence identified PSL22 as a Rho binding protein. Figure 15 shows a multiple sequence alignment of PSL22 with related proteins. Completely conserved residues are shown in black; residues found in three sequences are shaded.

The expression profile of *PSL22*, determined in various human tissues by Northern analysis (Figure 116B), revealed that while the highest expression was seen in the prostate, high expression was seen in the kidney, pancreas, and colon.

The androgen regulation of *PSL22* was examined in LNCaP cells, in the androgen-independent prostate cancer cell lines PC3 and DU145, and in CWR22R cells (Figure 16A). The results showed that *PSL22* is androgen regulated in LNCaP cells, where it is highly expressed, but is not androgen regulated in the PC3 and DU145 cells.

EXAMPLE 12

Materials And Methods

The following materials and methods were used in performing the exemplary experiments shown herein. It is understood that these materials and methods are subject to modifications that do not change the nature of the invention, as will be understood by those of ordinary skill in the art.

Probes

Poly (A)+ RNA	1 µg [(-) or (+)]
Random primer (N7)	200 ng

RNAse-free sterile H_2O to 20 ul

Heat at 70°C for 10 min, and chill on ice.

While heating the RNA samples, the following solution was prepared:

5X 1st strand buffer	10 µl
0.1 mM DTT	5 µl
10 mM each dTTP+dGTP	2 µl
^{32}P alpha dATP	5 µl
^{32}P alpha dCTP	5 µl
Superscript II (200 U /µl, BRL)	2µl

The solution was mixed by pipetting, spun briefly, incubated at 25°C for 5 min, and then for an additional 1 hour at 37°C. 2 µl of 10mM dCTP + dATP was added and the mixture was incubated for 30 min at 37°C and then heat inactivated at 70°C for 10 min. Unincorporated nucleotides were removed using prespun G25 columns (Bio-Rad). Specific activity (which should be over $5x10^8$cpm/µg) was calculated.

Hybridization

Freshly prepared 25 ml Hybridization mix (7% SDS, 0.5 M $NaHPO_4$, 1mM EDTA) was pre-warmed at 65°C and 12.5 ml was used for prehybridization of each membrane, 5-10 min at 65 °C. The probe was heat denatured at 95°C for 3-5 min and transferred to the prehybridization mix at 65°C. Hybridization was carried out at 65°C overnight.

Washing

Wash solution I (2xSSC and 1% SDS) and II (0.1xSSC and 0.5% SDS) were prewarmed, and the membrane were washed once with Solution I and then with Solution II for 30 min at 65°C. The membranes were covered with plastic wrap and exposed to a phoshorimager screen.

Selection

Clones that showed differences between the (-) and (+) blots were picked (usually 1-8 on each blot pair). A secondary round of reverse northern analysis for confirmation was performed, this time spotting each clone in duplicate on each blot. After phosphorimager analysis, the blots were stripped in 0.1xSSC and 0.5% SDS for 2x15 min at 95°C and hybridized with a PSA probe (or depending on the hormone that is being used, with a probe for any abundant target genes in the tissue under study). For the clones that were confirmed to be different from PSA, for differential expression in the secondary reverse northern, northern analysis was performed using established protocols. A time course of R1881 induction of LNCaP cells, as well as the CWR22 xenograft model upon androgen ablation (Wainstein, M. A. et al., *Cancer Res.* 54, 6049-6052, 1994) and the androgen-independent CWR22R relapsed xenograft (Nagabhushan, M. et al., *Cancer Res. 56,* 3042-6, 1996), was used.

Sequence analysis

Sequence analysis was perfomed by the dideoxy chain termination methods using an ABI automated sequencer. Homology search was done using a basic BLAST algorithm. Figure 3 shows a table of results obtained from the BLAST analysis of isolated clones and their homology to known genes (The cutoff for significant homology was 50% identity).

Isolation of prostate cancer related genes from LNCaP cells

The prostate cancer cell line LNCaP was cultured in two batches in culture conditions similar to those previously described (Horoszewicz JS et al., *Cancer Res.* 43: 1809-1818, 1983). The first batch was left untreated, while the second batch was treated with the synthetic androgen R1881 for 24 hrs. Cells from both batches were harvested and total RNA was then isolated from each batch. From the total RNA, polyA$^+$ RNA was obtained using standard procedures, and was used in the Suppression Subtraction Hybridization (SSH; Diatchenko et al., *supra*) procedure to identify hormone regulated genes. The tester in the SSH procedure was cDNA from untreated cells and the driver was cDNA from R1881-treated cells. The suppression subtraction protocol was performed according to the original description of the method (Diatchenko et al., *supra*).

Cell culture

LNCaP, PC-3 and DU-145 cells were routinely maintained and treated as described previously (Korkmaz, K. S. et al., *DNA Cell Biol* 19, 499-506, 2000; Korkmaz, K. S. et al., *Gene* 260, 25-36, 2000).

Xenograft studies

Transplantation, growth, and harvesting of tumors from mice bearing the CWR22 and CWR22R xenografts were as previously described (Wainstein, M. A., *supra*; Nagabhushan, M., *supra*).

Cloning and plasmid construction

A 262 bp cDNA fragment was originally obtained from a screen of a prostate specific library (Ausubel, F. M., et al. (1997) *Current Protocols in Molecular Biology* (John Wiley and Sons, New York) and termed L74. 5' Rapid Amplification of cDNA Ends (RACE) was performed (oligonucleotide sequences available upon request) using the Marathon-Ready cDNA that was prepared from normal prostate tissue (Clontech) and/or SMART-RACE LNCaP cDNA library (Clontech) that was generated according to the manufacturer's recommendations. RACE products were cloned into pCRII-TOPO (Invitrogen), positive clones were confirmed by Southern analysis, and sequenced. In parallel, a λgt10 cDNA library made from a pool of normal human prostates (Clontech) was screened by established procedures to obtain additional clones. Overlapping clones were used to deduce the full-length *STMP1* cDNA sequence.

The full-length *STMP1* ORF was amplified by using primers centered around the start and stop codons (sequences available upon request) and fused in frame to the C-terminus of green flourescent protein (GFP) using the vector pcDNA3.1-NT-GFP-TOPO (Invitrogen) to generate GFP-STMP1.

Northern analysis

Total RNA was prepared by the single step guanidine thiocyanate procedure and used in Northern analysis (18). 15 μg of total RNA was used per lane. Probes were generated by random priming and had a specific

activity of $>3x10^8$ dpm/μg. A cDNA fragment of *STMP1* spanning residues 145-2202 bp was used as probe. Bands were visualized and quantitated by phosphorimager analysis (Molecular Dynamics).

Confocal microscopy

COS-1 cells were transiently transfected by electroporation using a BTX square-wave pulser at 150 V, 1 ms duration. Cells were grown either on cover slips placed in 6-well tissue culture plates for indirect immunofluorescence or on Lab-Tek Chambered Coverglass (Nalge Nunc International) for live-cell microscopy. Transiently transfected cells were observed 16 h after transfection by Leica TCS-SP confocal microscope. All live-cell experiments were done at 37°C.

Indirect immunofluorescence

The indirect immunofluorescence was carried out as previously described (Misteli, T. & Spector, D. L. *Mol Cell* 3, 697-705, 1999). The following antibodies were used: anti-β-coat protein (β-COP) antiserum (kindly provided by J. Lippincott-Schwartz), anti-mannosidase II (kindly provided by T. Misteli), anti-TGN46 (Serotec, kindly provided by J.S. Bonifacino), and anti-EEA1 (Affinity Biotechnologies). Texas Red-conjugated secondary antibodies specific for mouse and rabbit were purchased from ICN Biomedicals (Costa Mesa, CA).

Other Embodiments

All publications and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each independent publication or patent application was specifically and individually indicated to be incorporated by reference.

While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and this application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure that come within known or customary practice within the art to which the invention pertains and may be applied to the essential features hereinbefore set forth, and follow in the scope of the appended claims.

CLAIMS

1. A substantially pure prostate-specific or testis-specific polypeptide, said polypeptide sequence comprising a sequence substantially identical to the sequence of any of SEQ ID NOS: 14, 29, 32, 34, 36, 41, or 53.

2. The substantially pure prostate-specific or testis-specific polypeptide of claim 1, said polypeptide sequence comprising the sequence of any of SEQ ID NOS: 14, 29, 32, 34, 36, 41, or 53.

3. An isolated nucleic acid molecule encoding a polypeptide of claim 1.

4. The isolated nucleic acid of claim 3, wherein said nucleic acid molecule comprises the sequence of any of SEQ ID NOS: 23, 28, 31, 33, 35, 40, or 52.

5. An isolated prostate-specific or testis-specific nucleic acid molecule, said nucleic acid molecule comprising a sequence substantially identical to SEQ ID NOS: 1-12, 22, 27, 30, and 51.

6. An isolated prostate-specific or testis-specific nucleic acid molecule, said nucleic acid molecule consisting essentially of SEQ ID NOS: 15-21, 24-26, 42-50, and 54-70.

7. The polypeptide of claim 1, wherein said polypeptide is derived from a mammal.

8. The polypeptide of claim 6, wherein said mammal is a human.

9. A vector comprising the isolated nucleic acid molecule of claim 3, 5, or 6.

10. A cell comprising the isolated nucleic acid molecule of claim 3, 5, or 6.

11. A cell comprising the vector of claim 9.

12. A non-human transgenic animal comprising the isolated nucleic acid molecule of claim 3, 5, or 6.

13. An isolated nucleic acid molecule that hybridizes under high stringency conditions to the complement of any of the sequences set forth in SEQ ID NOS: 1-12, 15-28, 30, 31, 33, 35, 40, 42-50, 51, 52, or 54-70, wherein said isolated nucleic acid molecule encodes a prostate-specific or testis-specific polypeptide.

14. An isolated nucleic acid molecule, wherein said nucleic acid molecule comprises a sequence that is antisense to the coding strand of any of the prostate-specific or testis-specific nucleic acid molecules set forth in SEQ ID NOS: 1-12, 15-28, 30, 31, 33, 35, 40, 42-50, 51, 52, or 54-70, or a fragment thereof.

15. A probe for analyzing a prostate-specific or testis-specific gene or homolog or fragment thereof, said probe having greater than 55% nucleotide sequence identity to a sequence encoding any of SEQ ID NOS: 1-12, 15-28, 30, 31, 33, 35, 40, 42-50, 51, 52, or 54-70, or fragment thereof,

wherein said fragment comprises at least six amino acids, and said probe hybridizes under high stringency conditions to at least a portion of a prostate-specific or testis-specific nucleic acid molecule.

16. The probe of claim 14, wherein said probe has 100% complementarity to a nucleic acid molecule encoding any of SEQ ID NOS: 1-12, 15-28, 30, 31, 33, 35, 40, 42-50, 51, 52, or 54-70, or fragment thereof,

wherein said fragment comprises at least six amino acids, and said probe hybridizes under high stringency conditions to at least a portion of a prostate-specific or testis-specific nucleic acid molecule.

17. An antibody that specifically binds to a prostate-specific or testis-specific polypeptide, said polypeptide comprising an amino acid sequence that is substantially identical to the amino acid sequence of any of SEQ ID NOS: 14, 29, 32, 34, 36, 41, or 53.

18. A method of detecting a prostate-specific or testis-specific gene or fragment thereof in a cell, said method comprising

contacting the nucleic acid molecule of any of SEQ ID NOS: 1-12, 15-28, 30, 31, 33, 35, 40, 42-50, 51, 52, or 54-70, or a fragment thereof, wherein said fragment is greater than about 18 nucleotides in length, with a preparation of genomic DNA from said cell, under high stringency hybridization conditions, and

detecting DNA sequences having about 55% or greater nucleotide sequence identity to any of SEQ ID NOS: 1-12, 15-28, 30, 31, 33, 35, 40, 42-50, 51, 52, or 54-70, thereby identifying a prostate-specific or testis-specific gene or fragment thereof.

19. A method for identifying a test compound that modulates the expression or activity of a prostate-specific or testis-specific polypeptide, said method comprising

contacting said prostate-specific or testis-specific polypeptide with said test compound, and

determining the effect of said test compound on said prostate-specific or testis-specific polypeptide expression or activity.

20. A method of treating a mammal having a disorder of the prostate or testis, said method comprising

administering to said mammal a therapeutically effective amount of a compound that modulates the activity or expression of a prostate-specific or testis-specific polypeptide,

wherein said compound has a beneficial effect on said disorder in said mammal.

21. The method of claim 20, wherein said disorder is prostate cancer.

22. A pharmaceutical composition comprising at least one dose of a therapeutically effective amount of a prostate-specific or testis-specific polypeptide or fragment thereof, in a pharmaceutically acceptable carrier, said composition being formulated for the treatment of a disorder of the prostate or testis.

23. The method of claim 20, wherein the prostate-specific or testis-specific polypeptide comprises an amino acid sequence substantially identical to the amino acid sequence of SEQ ID NOS: 14, 29, 32, 34, 36, 38, 39, 41, 53, or 71-73 and fragments and analogs thereof.

24. The method of claim 20, wherein said mammal is a human.

25. A kit for the analysis of a prostate-specific or testis-specific nucleic acid molecule, said kit comprising a nucleic acid molecule probe for analyzing a prostate-specific or testis-specific nucleic acid molecule present in a test subject.

26. A kit for the analysis of a prostate-specific or testis-specific polypeptide, said kit comprising an antibody for analyzing a prostate-specific or testis-specific polypeptide present in a test subject.

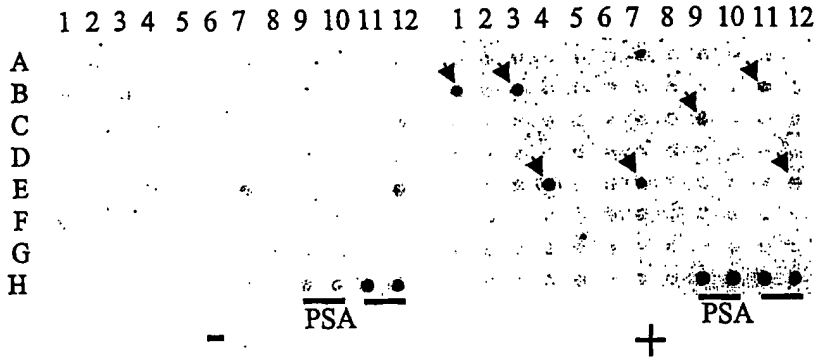
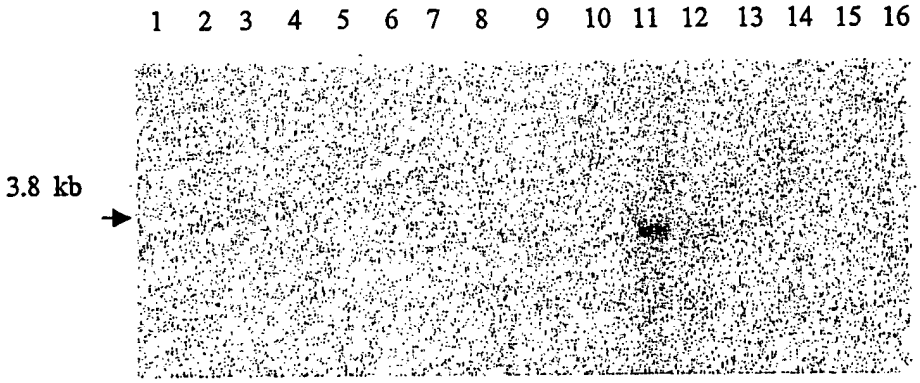


Figure 1



Multiple tissue northern blot. Lane 11: Prostate, 12:testis

Figure 2

SEQ. ID NO.	SEQ. NAME	LENGTH	SEQUENCE
1	PSL 22	251	ACTAATGTGAGGAAaCAAACATGTTCAGGCCTGAACATTTCCGGTGCTGACT CGGCcTTAAACGTTTGTGCCATAATGGAAAATATCTATCTATCTGTTCTCAA ATCCTGTTTTTCTCATAGTGTAAACTCACATTTGATGTGTTTTTATGAAGGAA AGTAACCAAGAAACCTCTAGGAATTAGGAAAAAAaGAACTTTTTTGAGGTG TGTTACTATACTGCTGTAAGTTATTTATTATATAAAGTATTGTAAATAGAAaT AGTGTTGAGATATGAAATATGGCTATTTTTAATGGTGACAATTATAGACTTT TAGgTCACTATTAAATTGGGGTTACCTATATCcAGT
2	PSL 229A	349	ACACATCCATCATTGTGAAATCTCTTTTCCAACAAACGTCCTCTTAATGAGC ACAATTCATTAAAaTCTTTGGGGACTAAGCTACGAACAAAGTTCAACTAAAC TACCTACTGACTTCAAAAGGAACATATACCCACCACGTGTGGTAGCTCATG ACTGTAATCCCAGCACTTTGGGAGGCTGAGGCAGGAGGATCACCTGAGCCC AGGAGTTCCAGACCAGCCTAAGCAACATGCCAAGACCCTGTATGT
3	PSL E15C	51	ACAAAGACACCCTTGTYCCCCGGGCAAGGTCCTCCAGCTACAAGGGGGCCA
4	PSL E156	148	CCXYACATTGTCACAGAGAGGCTCCAGGCTTAAAGTTGACCTGCGTAGAAA GCAAGAATGAATTGTTGGAGGAAGTAAGGAGGGCGATTGAATAAAGACTTT TAGCAGCTGGGCCAGCTGAACCATCCCAACCCTTCAAATCCCCTTGT
5	PSL E157	261	ACCCTAACTGAACCCATTTCAGCCACTCAGATTGATAGGGTGGAAAAGACA GGGCAGGTGGTAGCAGCTGTGAAGAAAAGAGGAAAGCAGAAGGGTGGCCT ATAATCTACAGGCATGTAGAGAGGACTACATAGGCCTCTGTTCTTTGCCCTC AGGAGCCCCCTTCCTGTCCCTTGGACTCAGAATGGATCCTTCCAGCACACAT GGCCCAACACTGAGAGTGCAGGAAGCATGGGTAGGGGCCTCCTGCTGCTGG TATGT
6	PSL E391	121	AGTNTGNGGGGANTTGAGGGCNGNTACGNNAAANGNTGGNCTACTNTAGA TGCTGCTCGAGCGGCCGCCAGTGTGATGGATACAAGCTTTCTTTTTTTTTTTT ATTTTCGNNTTTTTTTC
7	PSL K31	93	ACTCAGTAGGGACTGAGCACTAAATGCTTATTTTAAAAGAAATGTAAAGAG CAGAAAGCAATTCAGGCTACCCTGCCTTTTGTGCTGGCTAGT
8	PSL L28	169	ACACTTAAAATAGTTAATGTGATACATTTTATGTTACATGTATTTTGCCCAC TGAAAAAATAAAAATATATAAACACACAGCAAATGATGACCAGGCCTTTGA AGAAAGCTTATAAAACAAAATTAAGAAGCCTGGCTACAGAGCGAGACTCTG TCTCAAAAAAAAAAA
9	PSL L74	262	ACTTTACAAGCATGAAGGATATTAGGGTAAGTGGCTAATTATAAATCTACT CTAGAGACATATAATCATACAGATTATTCATAAAATTTTTCAGTGCTGTCCT TCCACATTTAATTGCATTTTGCTCAAACTGTAGAATGCCCTACATTCCCCCC ACCCCAATTTGCTATTTCCTTATTAAAATAGAAAATTATAGGCAAGATACAA TTATATGCGTTCCTCTTCCTGAAATTATAACATTTCTAAACTTACCCACGTAG GT
10	PSL SSH 20	175	ACAGGTTGGCCCTTCACCTAGTTGACTCAGCCCTCGATAGTCTAGAGCCCAC CCCCTCCTCAGGAACTCAAGAGCTCAGCATTTATAATGAGCAGTTGGTAAT GAGTTGCCCTATGTGCTTGTCGCAAGCAGTCACAGAGATGAGCCCTATTACT TGATATTCAGGAACAAAGGT
11	PSL SSH 4	331	ACATCCAAGCCTTCCTCTGCGTGAGAGCAAAGGCTTTGCTCATCAGCCAGCC AGTCTTGTTACTATCTGGCTACTTTTTAAGGTTAAAAAATAAAAGGCAGTTT CTTTGCTCTGCAGGCGGCAAGGCAGGAGGCGCAGGCCTCTTCATTGTTCAC ATGTCACAGGAGGAGGCTCTGAGCAAAGGCCACTGGCAAGTTAGGGCAAC ACCAAGAAGGCTCTGCGGAGAGACTCCCTGTGGGTTGGGGGsCTGGCAGGA ACGGTGCcTGTGGACTGTTTATGGTCTGTCCAGTTGAGGCTTGGTAAACCCA AGTAAAGTGTTAAAAACCTCAGT
12	PSL SSH 9	170	ACGACTCATCCACCTCCGGCTGAAGCTCCAGGAGCTGAAGGACCCCAATGA GGATGAGCCAAACATCCGAGTGCTCCTTGAGCACCGCTTTTACAAGGAGAA GAGCAAGAGCGTCAAGCAGACCTGTGACAAGTGTAACACCATCATCTGGGG GCTCATTCAGACCTGGT

Figure 3

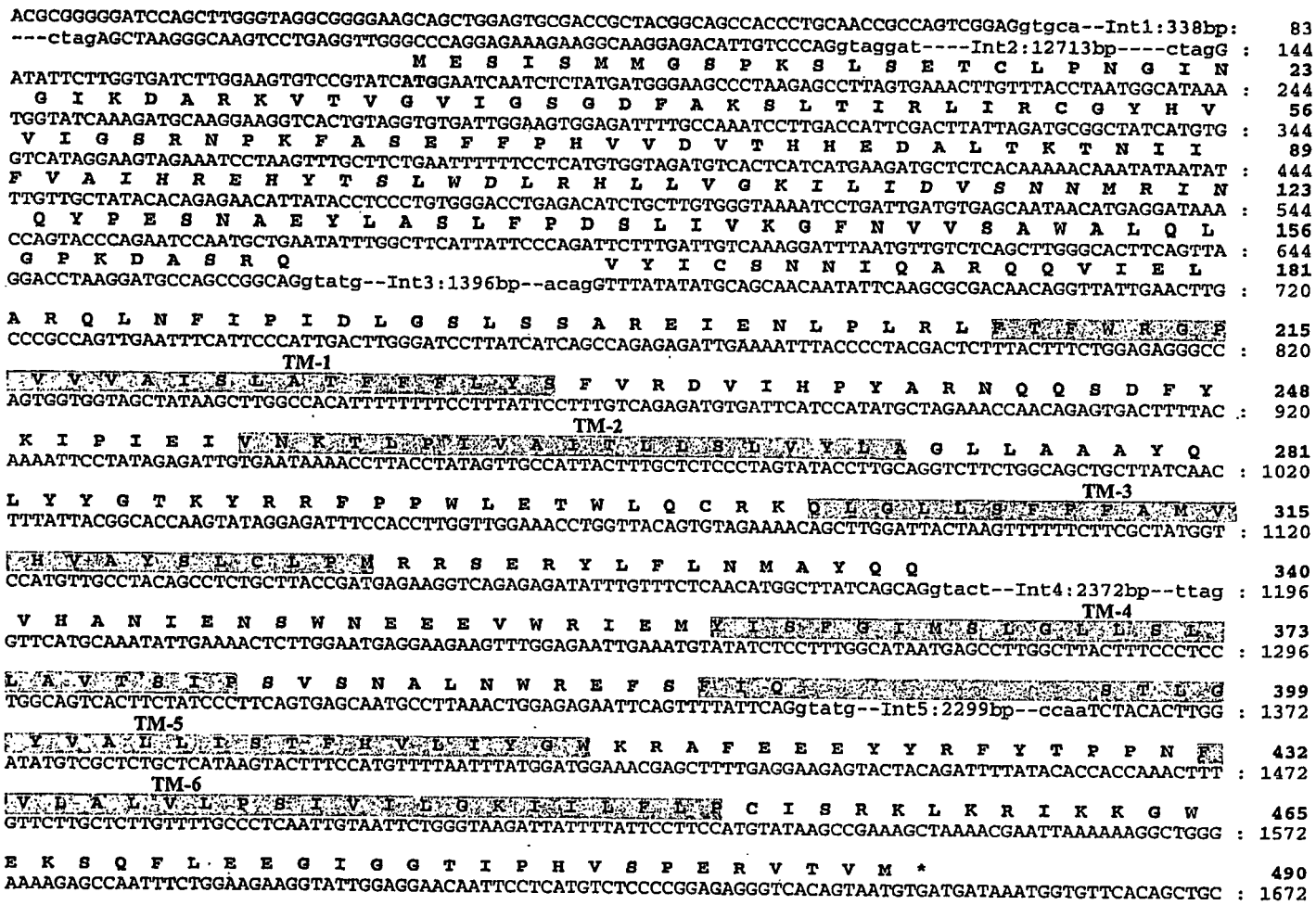


Figure 4B

```
EXON_1  83bp
     1  ACGCGGGGGATCCAGCTTGGGTAGGCGGGGAAGCAGCTGGAGTGCGACCGCTACGGCAGC
    61  CACCCTGCAACCGCCAGTCGGAG

EXON_2  61bp
     1  AGCTAAGGGCAAGTCCTGAGGTTGGGCCCAGGAGAAAGAAGGCAAGGAGACATTGTCCCA
    61  G

EXON_3  525bp
     1  GATATTCTTGGTGATCTTGGAAGTGTCCGTATCATGGAATCAATCTCTATGATGGGAAGC
    61  CCTAAGAGCCTTAGTGAAACTTTTTTACCTAATGGCATAAATGGTATCAAAGATGCAAGG
   121  AAGGTCACTGTAGGTGTGATTGGAAGTGGAGATTTTGCCAAATCCTTGACCATTCGACTT
   181  ATTAGATGCGGCTATCATGTGGTCATAGGAAGTAGAAATCCTAAGTTTGCTTCTGAATTT
   241  TTTCCTCATGTGGTAGATGTCACTCATCATGAAGATGCTCTCACAAAAACAAATATAATA
   301  TTTGTTGCTATACACAGAGAACATTATACCTCCCTGTGGGACCTGAGACATCTGCTTGTG
   361  GGTAAAATCCTGATTGATGTGAGCAATAACATGAGGATAAACCAGTACCCAGAATCCAAT
   421  GCTGAATATTTGGCTTCATTATTCCCAGATTCTTTGATTGTCAAAGGATTTAATGTTGTC
   481  TCAGCTTGGGCACTTCAGTTAGGACCTAAGGATGCCAGCCGGCAG

EXON_4  528bp
     1  GTTTATATATGCAGCAACAATATTCAAGCGCGACAACAGGTTATTGAACTTGCCCGCCAG
    61  TTGAATTTCATTCCCATTGACTTGGGATCCTTATCATCAGCCAGAGAGATTGAAAATTTA
   121  CCCCTACGACTCTTTACTCTCTGGAGAGGGCCAGTGGTGGTAGCTATAAGCTTGGCCACA
   181  TTTTTTTTCCTTTATTCCTTTGTCAGAGATGTGATTCATCCATATGCTAGAAACCAACAG
   241  AGTGACTTTTACAAAATTCCTATAGAGATTGTGAATAAAACCTTACCTATAGTTGCCATT
   301  ACTTTGCTCTCCCTAGTATACCTCGCAGGTCTTCTGGCAGCTGCTTATCAACTTTATTAC
   361  GGCACCAAGTATAGGAGATTTCCACCTTGGTTGGAAACCTGGTTACAGTGTAGAAAACAG
   421  CTTGGATTACTAAGTTTTTTCTTCGCTATGGTCCATGTTGCCTACAGCCTCTGCTTACCG
   481  ATGAGAAGGTCAGAGAGATATTTGTTTCTCAACATGGCTTATCAGCAG

EXON_5  165bp
     1  GTTCATGCAAATATTGAAAACTCTTGGAATGAGGAAGAAGTTTGGAGAATTGAAATGTAT
    61  ATCTCCTTTGGCATAATGAGCCTTGGCTTACTTTCCCTCCTGGCAGTCACTTCTATCCCT
   121  TCAGTGAGCAATGCTTTAAACTGGAGAGAATTCAGTTTTATTCAG

EXON_6  148bp
     1  TCTACACTTGGATATGTCGCTCTGCTCATAAGTACTTTCCATGTTTTAATTTATGGATGG
    61  AAACGAGCTTTTGAGGAAGAGTACTACAGATTTTATACACCACCAAACTTTGTTCTTGCT
   121  CTTGTTTTGCCCTCAATTGTAATTCTGG

cont_6+UTR
     1  GTAAGATTATTTTATTCCTTCCATGTATAAGCCGAAAGCTAAAACGAATTAAAAAAGGCT
    61  GGGAAAAGAGCCAATTTCTGGAAGAAGGTATTGGAGGAACAATTCCTCATGTCTCCCCGG
   121  AGAGGGTCACAGTAATGTGATGATAAATGGTGTTCACAGCTGCCATATAAAGTTCTACTC
   181  ATGCCATTATTTTTATGACTTCTACGTTCAGTTACAAGTATGCTGTCAAATTATCGTGGG
   241  TTGAAACTTGTTAAATGAGATTTCAACTGACTTAGTGATAGAGTTTTCTTCAAGTTAATT
   301  TTCACAAATGTCATGTTTGCCAATATGAATTTTTCTAGTCAACATATTATTGTAATTTAG
   361  GTATGTTTTGTTTTGTTTTGCACAACTGTAACCCTGTTGTTACTTTATATTTCATAATCA
   421  GACAAAAATACTTACAGTTAATAATATAGATATAATGTTAAAAACAATTTGCAAACCAGC
   481  AGAATTTTAAGCTTTTAAAATAATTCAATGGATATACATTTTTTTCTGAAGATTAAGATT
   541  TTAATTATTCAACTTAAAAAGTAGAAATGCATTATTATACATTTTTTTAAGAAAGGACAC
   601  GTTATGTTAGCATCTAGGTAAGGCTGCATGATAGCATTCCTATATTTCTCTCATAAAATA
   661  GGATTTGAAGGATGAAATTAATTGTATGAAGCAATGTGATTATATGAAGAGACACAAATT
   721  AAAAAGACAAATTAAACCTGAAATTATATTTAAAATATATTTGAGACATGAAATACATAC
   781  TGATAATACATACCTCATGAAAGATTTTATTCTTTATTGTGTTACAGAGCAGTTTCATTT
   841  TCATATTAATATACTGATCAGGAAGAGGATTCAGTAACATTTGGCTTCCAAAACTGCTAT
   901  CTCTAATACGGTACCAATCCTAGGAACTGTATACTAGTTCCTACTTAGAACAAAAGTATC
```

FIGURE 4C

```
 961  AAGTTTGCACACAAGTAATCTGCCAGCTGACCTTTGTCGCACCTTAACCAGTCACCACTT
1021  GCTATGGTATAGGATTATACTGATGTTCTTTGAGGGATTCTGATGTGCTAGGCATGGTTC
1081  TAAGTACTTTACTTGTATTATCCCATTTAATACTTAGAACAACCCCGTGAGATAAGTAGT
1141  TATTATCCTCATTTTACACATGAGGGACCGAAGGATAGAAAAGTTATTTTTCAAAGGTCT
1201  TGCAGTTAATAAATGGCAGAGTGAGCATTCAAGTCCAGGTAGTCATATTCCAGAGGCCAC
1261  GGTTTTAACCACTAGGCTCTAGAGCTCCCGCCGCGCCCCTATGCATTATGTTCACAATGC
1321  CAATCTAGATGCTTCCTCTTTTGTATAAAGTCACTGACATTCTTTAGAGTGGGTTGGGTG
1381  CATCCAAAAATGTATAAAAATATTATTATAATAAACTTATTACTGCTTGTAGGGTAATTC
1441  ACAGTTACTTACCCTATTCTTGCTTGGAACATGAGCCTGGAGACCCATGGCAGTCCATAT
1501  GCCTCCCTATGCAGTGAAGGGCCCTAGCAGTGTTAACAAATTGCTGAGATCCCACGGAGT
1561  CTTTCAAAAATCTCTGTAGAGTTAGTCTTCTCCTTTTCTCTTCCTGAGAAGTTCTCCTGC
1621  CTGCATAACCATTCATTAGGGAGTACTTTACAAGCATGAAGGATATTAGGGTAAGTGGCT
1681  AATTATAAATCTACTCTAGAGACATATAATCATACAGATTATTCATAAAATTTTTCAGTG
1741  CTGTCCTTCCACATTTAATTGCATTTTGCTCAAACTGTAGAATGCCCTACATTCCCCCCA
1801  CCCCAATTTGCTATTTCCTTATTAAAATAGAAAATTATAGGCAAGATACAATTATATGCG
1861  TTCCTCTTCCTGAAATTATAACATTTCTAAACTTACCCACGTAGGGACTACTGAATCCAA
1921  CTGCCAACAATAAAAAGACTTTTATTTAGTAGAGGCTACCTTTCCCCCCAGTGACTCTTT
1981  TTCTACAACTGCCTTGTCAGTTTGGTAATTCACTTATGATTTTCTAATGTTCTCTTGGTG
2041  AATTTTATTATCTTGGACCCTCTTTTTTTTTTTTTTTTAAAGACAGAGTCTTGCTCTGTCA
2101  CCCATTGCTCTCGTTTGGGCAACAAGAGTGAAACTCTTGTCTCAAAAAAAAAAAAAAATG
2161  AGGTTTAAGACAGTTTTGTCATTACTGGTGGGATCTGGTCACACAAGATAGCATTAAACG
2221  TGACATGGCACATAAAATTGGTTAAAAAATTTTGTTTTTTAATTGCGTAATGTAAAAGCC
2281  CAACAAACACTTTATGCAAGATTGGAATGTATCTTCAAATTCAGATTTAATAAACATGTA
2341  AAGATCCTCTGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 4C

```
   1  ACGCGGGGGATCCAGCTTGGGTAGGCGGGGAAGCAGCTGGAGTGCGACCGCTACGGCAGC
  61  CACCCTGCAACCGCCAGTCGGAGAGCTAAGGGCAAGTCCTGAGGTTGGGCCCAGGAGAAA
 121  GAAGGCAAGGAGACATTGTCCCAGGATATTCTTGGTGATCTTGGAAGTGTCCGTATCATG
 181  GAATCAATCTCTATGATGGGAAGCCCTAAGAGCCTTAGTGAAACTTGTTTACCTAATGGC
 241  ATAAATGGTATCAAAGATGCAAGGAAGGTCACTGTAGGTGTGATTGGAAGTGGAGATTTT
 301  GCCAAATCCTTGACCATTCGACTTATTAGATGCGGCTATCATGTGGTCATAGGAAGTAGA
 361  AATCCTAAGTTTGCTTCTGAATTTTTTCCTCATGTGGTAGATGTCACTCATCATGAAGAT
 421  GCTCTCACAAAAACAAATATAATATTTGTTGCTATACACAGAGAACATTATACCTCCCTG
 481  TGGGACCTGAGACATCTGCTTGTGGGTAAAATCCTGATTGATGTGAGCAATAACATGAGG
 541  ATAAACCAGTACCCAGAATCCAATGCTGAATATTTGGCTTCATTATTCCCAGATTCTTTG
 601  ATTGTCAAAGGATTTAATGTTGTCTCAGCTTGGGCACTTCAGTTAGGACCTAAGGATGCC
 661  AGCCGGCAGGTTTATATATGCAGCAACAATATTCAAGCGCGACAACAGGTTATTGAACTT
 721  GCCCGCCAGTTGAATTTCATTCCCATTGACTTGGGATCCTTATCATCAGCCAGAGAGATT
 781  GAAAATTTACCCCTACGACTCTTTACTTTCTGGAGAGGGCCAGTGGTGGTAGCTATAAGC
 841  TTGGCCACATTTTTTTTCCTTTATTCCTTTGTCAGAGATGTGATTCATCCATATGCTAGA
 901  AACCAACAGAGTGACTTTTACAAAATTCCTATAGAGATTGTGAATAAAACCTTACCTATA
 961  GTTGCCATTACTTTGCTCTCCCTAGTATACCTTGCAGGTCTTCTGGCAGCTGCTTATCAA
1021  CTTTATTACGGCACCAAGTATAGGAGATTTCCACCTTGGTTGGAAACCTGGTTACAGTGT
1081  AGAAAACAGCTTGGATTACTAAGTTTTTTCTTCGCTATGGTCCATGTTGCCTACAGCCTC
1141  TGCTTACCGATGAGAAGGTCAGAGAGATATTTGTTTCTCAACATGGCTTATCAGCAGGTT
1201  CATGCAAATATTGAAAACTCTTGGAATGAGGAAGAAGTTTGGAGAATTGAAATGTATATC
1261  TCCTTTGGCATAATGAGCCTTGGCTTACTTTCCCTCCTGGCAGTCACTTCTATCCCTTCA
1321  GTGAGCAATGCCTTAAACTGGAGAGAATTCAGTTTTATTCAGTCTACACTTGGATATGTC
1381  GCTCTGCTCATAAGTACTTTCCATGTTTTAATTTATGGATGGAAACGAGCTTTTGAGGAA
1441  GAGTACTACAGATTTTATACACCACCAAACTTTGTTCTTGCTCTTGTTTTGCCCTCAATT
1501  GTAATTCTGGGTAAGATTATTTTATTCCTTCCATGTATAAGCCGAAAGCTAAAACGAATT
1561  AAAAAAGGCTGGGAAAAGAGCCAATTTCTGGAAGAAGGTATTGGAGGAACAATTCCTCAT
1621  GTCTCCCCGGAGAGGGTCACAGTAATGTGATGATAAATGGTGTTCACAGCTGCCATATAA
1681  AGTTCTACTCATGCCATTATTTTTATGACTTCTACGTTCAGTTACAAGTATGCTGTCAAA
1741  TTATCGTGGGTTGAAACTTGTTAAATGAGATTTCAACTGACTTAGTGATAGAGTTTTCTT
1801  CAAGTTAATTTTCACAAATGTCATGTTTGCCAATATGAATTTTTCTAGTCAACATATTAT
1861  TGTAATTTAGGTATGTTTTGTTTTGTTTTGCACAACTGTAACCCTGTTGTTACTTTATAT
1921  TTCATAATCAGACAAAAATACTTACAGTTAATAATATAGATATAATGTTAAAAACAATTT
1981  GCAAACCAGCAGAATTTTAAGCTTTTAAAATAATTCAATGGATATACATTTTTTTCTGAA
2041  GATTAAGATTTTAATTATTCAACTTAAAAAGTAGAAATGCATTATTATACATTTTTTTAA
2101  GAAAGGACACGTTATGTTAGCATCTAGGTAAGGCTGCATGATAGCATTCCTATATTTCTC
2161  TCATAAAATAGGATTTGAAGGATGAAATTAATTGTATGAAGCAATGTGATTATATGAAGA
2221  GACACAAATTAAAAAGACAAATTAAACCTGAAATTATATTTAAAATATATTTGAGACATG
2281  AAATACATACTGATAATACATACCTCATGAAAGATTTTATTCTTTATTGTGTTACAGAGC
2341  AGTTTCATTTTCATATTAATATACTGATCAGGAAGAGGATTCAGTAACATTTGGCTTCCA
2401  AAACTGCTATCTCTAATACGGTACCAATCCTAGGAACTGTATACTAGTTCCTACTTAGAA
2461  CAAAAGTATCAAGTTTGCACACAAGTAATCTGCCAGCTGACCTTTGTCGCACCTTAACCA
2521  GTCACCACTTGCTATGGTATAGGATTATACTGATGTTCTTTGAGGGATTCTGATGTGCTA
2581  GGCATGGTTCTAAGTACTTTACTTGTATTATCCCATTTAATACTTAGAACAACCCCGTGA
2641  GATAAGTAGTTATTATCCTCATTTTACACATGAGGGACCGAAGGATAGAAAAGTTATTTT
2701  TCAAAGGTCTTGCAGTTAATAAATGGCAGAGTGAGCATTCAAGTCCAGGTAGTCATATTC
2761  CAGAGGCCACGGTTTTAACCACTAGGCTCTAGAGCTCCCGCCGCGCCCCTATGCATTATG
2821  TTCACAATGCCAATCTAGATGCTTCCTCTTTTGTATAAAGTCACTGACATTCTTTAGAGT
2881  GGGTTGGGTGCATCCAAAAATGTATAAAAATATTATTATAATAAACTTATTACTGCTTGT
2941  AGGGTAATTCACAGTTACTTACCCTATTCTTGCTTGGAACATGAGCCTGGAGACCCATGG
3001  CAGTCCATATGCCTCCCTATGCAGTGAAGGGCCCTAGCAGTGTTAACAAATTGCTGAGAT
3061  CCCACGGAGTCTTTCAAAAATCTCTGTAGAGTTAGTCTTCTCCTTTTCTCTTCCTGAGAA
3121  GTTCTCCTGCCTGCATAACCATTCATTAGGGAGTACTTTACAAGCATGAAGGATATTAGG
3181  GTAAGTGGCTAATTATAAATCTACTCTAGAGACATATAATCATACAGATTATTCATAAAA
3241  TTTTTCAGTGCTGTCCTTCCACATTTAATTGCATTTTGCTCAAACTGTAGAATGCCCTAC
3301  ATTCCCCCCACCCCAATTTGCTATTTCCTTATTAAAAATGTATAAAAATATTATTATAAT
3361  AAACTTATTACTGCTTGTAGGGTAATTCACAGTTACTTACCCTATTCTTGCTTGGAACAT
```

FIGURE 4D

```
3421  GAGCCTGGAGACCCATGGCAGTCCATATGCCTCCCTATGCAGTGAAGGGCCCTAGCAGTG
3481  TTAACAAATTGCTGAGATCCCACGGAGTCTTTCAAAAATCTCTGTAGAGTTAGTCTTCTC
3541  CTTTTCTCTTCCTGAGAAGTTCTCCTGCCTGCATAACCATTCATTAGGGAGTACTTTACA
3601  AGCATGAAGGATATTAGGGTAAGTGGCTAATTATAAATCTACTCTAGAGACATATAATCA
3661  TACAGATTATTCATAAAATTTTTCAGTGCTGTCCTTCCACATTTAATTGCATTTTGCTCA
3721  AACTGTAGAATGCCCTACATTCCCCCCACCCCAATTTGCTATTTCCTTATTAAAATAGAA
3781  AATTATAGGCAAGATACAATTATATGCGTTCCTCTTCCTGAAATTATAACATTTCTAAAC
3841  TTACCCACGTAGGGACTACTGAATCCAACTGCCAACAATAAAAAGACTTTTATTTAGTAG
3901  AGGCTACCTTTCCCCCCAGTGACTCTTTTTCTACAACTGCCTTGTCAGTTTGGTAATTCA
3961  CTTATGATTTTCTAATGTTCTCTTGGTGAATTTTATTATCTTGGACCCTCTTTTTTTTTT
4021  TTTTTAAAGACAGAGTCTTGCTCTGTCACCCATTGCTCTCGTTTGGGCAACAAGAGTGAA
4081  ACTCTTGTCTCAAAAAAAAAAAAAAATGAGGTTTAAGACAGTTTTGTCATTACTGGTGGG
4141  ATCTGGTCACACAAGATAGCATTAAACGTGACATGGCACATAAAATTGGTTAAAAAATTT
4201  TGTTTTTTAATTGCGTAATGTAAAAGCCCAACAAACACTTTATGCAAGATTGGAATGTAT
4261  CTTCAAATTCAGATTTAATAAACATGTAAAGATCCTCTGTAAAAAAAAAAAAAAAAAAA
4321  AAAAAAAAA
```

FIGURE 4D

```
   1  ACGCGGGGGATCCAGCTTGGGTAGGCGGGGAAGCAGCTGGAGTGCGACCGCTACGGCAGC
  61  CACCCTGCAACCGCCAGTCGGAGAGCTAAGGGCAAGTCCTGAGGTTGGGCCCAGGAGAAA

   1                                                           M
 121  GAAGGCAAGGAGACATTGTCCCAGGATATTCTTGGTGATCTTGGAAGTGTCCGTATCATG

   2   E  S  I  S  M  M  G  S  P  K  S  L  S  E  T  C  L  P  N  G
 181  GAATCAATCTCTATGATGGGAAGCCCTAAGAGCCTTAGTGAAACTTGTTTACCTAATGGC

  22   I  N  G  I  K  D  A  R  K  V  T  V  G  V  I  G  S  G  D  F
 241  ATAAATGGTATCAAAGATGCAAGGAAGGTCACTGTAGGTGTGATTGGAAGTGGAGATTTT

  42   A  K  S  L  T  I  R  L  I  R  C  G  Y  H  V  V  I  G  S  R
 301  GCCAAATCCTTGACCATTCGACTTATTAGATGCGGCTATCATGTGGTCATAGGAAGTAGA

  62   N  P  K  F  A  S  E  F  F  P  H  V  V  D  V  T  H  H  E  D
 361  AATCCTAAGTTTGCTTCTGAATTTTTTCCTCATGTGGTAGATGTCACTCATCATGAAGAT

  82   A  L  T  K  T  N  I  I  F  V  A  I  H  R  E  H  Y  T  S  L
 421  GCTCTCACAAAAACAAATATAATATTTGTTGCTATACACAGAGAACATTATACCTCCCTG

 102   W  D  L  R  H  L  L  V  G  K  I  L  I  D  V  S  N  N  M  R
 481  TGGGACCTGAGACATCTGCTTGTGGGTAAAATCCTGATTGATGTGAGCAATAACATGAGG

 122   I  N  Q  Y  P  E  S  N  A  E  Y  L  A  S  L  F  P  D  S  L
 541  ATAAACCAGTACCCAGAATCCAATGCTGAATATTTGGCTTCATTATTCCCAGATTCTTTG

 142   I  V  K  G  F  N  V  V  S  A  W  A  L  Q  L  G  P  K  D  A
 601  ATTGTCAAAGGATTTAATGTTGTCTCAGCTTGGGCACTTCAGTTAGGACCTAAGGATGCC

 162   S  R  Q  V  Y  I  C  S  N  N  I  Q  A  R  Q  Q  V  I  E  L
 661  AGCCGGCAGGTTTATATATGCAGCAACAATATTCAAGCGCGACAACAGGTTATTGAACTT

 182   A  R  Q  L  N  F  I  P  I  D  L  G  S  L  S  S  A  R  E  I
 721  GCCCGCCAGTTGAATTTCATTCCCATTGACTTGGGATCCTTATCATCAGCCAGAGAGATT

 202   E  N  L  P  L  R  L  F  T  F  W  R  G  P  V  V  V  A  I  S
 781  GAAAATTTACCCCTACGACTCTTTACTTTCTGGAGAGGGCCAGTGGTGGTAGCTATAAGC

 222   L  A  T  F  F  F  L  Y  S  F  V  R  D  V  I  H  P  Y  A  R
 841  TTGGCCACATTTTTTTTCCTTTATTCCTTTGTCAGAGATGTGATTCATCCATATGCTAGA

 242   N  Q  Q  S  D  F  Y  K  I  P  I  E  I  V  N  K  T  L  P  I
 901  AACCAACAGAGTGACTTTTACAAAATTCCTATAGAGATTGTGAATAAAACCTTACCTATA

 262   V  A  I  T  L  L  S  L  V  Y  L  A  G  L  L  A  A  A  Y  Q
 961  GTTGCCATTACTTTGCTCTCCCTAGTATACCTTGCAGGTCTTCTGGCAGCTGCTTATCAA

 282   L  Y  Y  G  T  K  Y  R  R  F  P  P  W  L  E  T  W  L  Q  C
1021  CTTTATTACGGCACCAAGTATAGGAGATTTCCACCTTGGTTGGAAACCTGGTTACAGTGT

 302   R  K  Q  L  G  L  L  S  F  F  F  A  M  V  H  V  A  Y  S  L
1081  AGAAAACAGCTTGGATTACTAAGTTTTTTCTTCGCTATGGTCCATGTTGCCTACAGCCTC

 322   C  L  P  M  R  R  S  E  R  Y  L  F  L  N  M  A  Y  Q  Q  V
1141  TGCTTACCGATGAGAAGGTCAGAGAGATATTTGTTTCTCAACATGGCTTATCAGCAGGTT
```

FIGURE 4E

```
 342   H  A  N  I  E  N  S  W  N  E  E  E  V  W  R  I  E  M  Y  I
1201   CATGCAAATATTGAAAACTCTTGGAATGAGGAAGAAGTTTGGAGAATTGAAATGTATATC

 362   S  F  G  I  M  S  L  G  L  L  S  L  L  A  V  T  S  I  P  S
1261   TCCTTTGGCATAATGAGCCTTGGCTTACTTTCCCTCCTGGCAGTCACTTCTATCCCTTCA

 382   V  S  N  A  L  N  W  R  E  F  S  F  I  Q  S  T  L  G  Y  V
1321   GTGAGCAATGCCTTAAACTGGAGAGAATTCAGTTTTATTCAGTCTACACTTGGATATGTC

 402   A  L  L  I  S  T  F  H  V  L  I  Y  G  W  K  R  A  F  E  E
1381   GCTCTGCTCATAAGTACTTTCCATGTTTTAATTTATGGATGGAAACGAGCTTTTGAGGAA

 422   E  Y  Y  R  F  Y  T  P  P  N  F  V  L  A  L  V  L  P  S  I
1441   GAGTACTACAGATTTTATACACCACCAAACTTTGTTCTTGCTCTTGTTTTGCCCTCAATT

 442   V  I  L  G  K  I  I  L  F  L  P  C  I  S  R  K  L  K  R  I
1501   GTAATTCTGGGTAAGATTATTTTATTCCTTCCATGTATAAGCCGAAAGCTAAAACGAATT

 462   K  K  G  W  E  K  S  Q  F  L  E  E  G  I  G  G  T  I  P  H
1561   AAAAAAGGCTGGGAAAAGAGCCAATTTCTGGAAGAAGGTATTGGAGGAACAATTCCTCAT

 482   V  S  P  E  R  V  T  V  M  *
1621   GTCTCCCCGGAGAGGGTCACAGTAATGTGATGATAAATGGTGTTCACAGCTGCCATATAA
```

FIGURE 4E

```
EXON_1  75bp
     1  GATCCAGCTTGGGTAGGCGGGGAAGCAGCTGGAGTGCGACCGCCGCGGCAGCCACCCTGC
    61  AACCGCCAGTCGGAG

EXON_2  79bp
     1  AGAGCTAAGGGCAAGTCCTGAGGTTGGGCCCAGGAGAAAGAAGGCAAGGAGACATTGTCC
    61  CAGGTAGGATGTGTCCCAG

EXON_3  525bp
     1  GATATTCTTGGTGATCTTGGAAGTGTCCGTATCATGGAATCAATCTCTATGATGGGAAGC
    61  CCTAAGAGCCTTAGTGAAACTTTTTTACCTAATGGCATAAATGGTATCAAAGATGCAAGG
   121  AAGGTCACTGTAGGTGTGATTGGAAGTGGAGATTTTGCCAAATCCTTGACCATTCGACTT
   181  ATTAGATGCGGCTATCATGTGGTCATAGGAAGTAGAAATCCTAAGTTTGCTTCTGAATTT
   241  TTTCCTCATGTGGTAGATGTCACTCATCATGAAGATGCTCTCACAAAAACAAATATAATA
   301  TTTGTTGCTATACACAGAGAACATTATACCTCCCTGTGGGACCTGAGACATCTGCTTGTG
   361  GGTAAAATCCTGATTGATGTGAGCAATAACATCAGGATAAACCAGTACCCAGAATCCAAT
   421  GCTGAATATTTGGCTTCATTATTCCCAGATTCTTTGATTGTCAAAGGATTTAATGTTGTC
   481  TCAGCTTGGGCACTTCAGTTAGGACCTAAGGATGCCAGCCGGCAG

EXON_4  528bp
     1  GTTTATATATGCAGCAACAATATTCAAGCGCGACAACAGGTTATTGAACTTGCCCGCCAG
    61  TTGAATTTCATTCCCATTGACTTGGGATCCTTATCATCAGCCAGAGAGATTGAAAATTTA
   121  CCCCTACGACTCTTTACTCTCTGGAGAGGGCCAGTGGTGGTAGCTATAAGCTTGGCCACA
   181  TTTTTTTTCCTTTATTCCTTTGTCAGAGATGTGATTCATCCATATGCTAGAAACCAACAG
   241  AGTGACTTTTACAAAATTCCTATAGAGATTGTGAATAAAACCTTACCTATAGTTGCCATT
   301  ACTTTGCTCTCCCTAGTATACCTCGCAGGTCTTCTGGCAGCTGCTTATCAACTTTATTAC
   361  GGCACCAAGTATAGGAGATTTCCACCTTGGTTGGAAACCTGGTTACAGTGTAGAAAACAG
   421  CTTGGATTACTAAGTTTTTTCTTCGCTATGGTCCATGTTGCCTACAGCCTCTGCTTACCG
   481  ATGAGAAGGTCAGAGAGATATTTGTTTCTCAACATGGCTTATCAGCAG

EXON_5  165bp
     1  GTTCATGCAAATATTGAAAACTCTTGGAATGAGGAAGAAGTTTGGAGAATTGAAATGTAT
    61  ATCTCCTTTGGCATAATGAGCCTTGGCTTACTTTCCCTCCTGGCAGTCACTTCTATCCCT
   121  TCAGTGAGCAATGCTTTAAACTGGAGAGAATTCAGTTTTATTCAG

EXON_6 148bp
     1  TCTACACTTGGATATGTCGCTCTGCTCATAAGTACTTTCCATGTTTTAATTTATGGATGG
    61  AAACGAGCTTTTGAGGAAGAGTACTACAGATTTTATACACCACCAAACTTTGTTCTTGCT
   121  CTTGTTTTGCCCTCAATTGTAATTCTGG

EXON_7+UTR  718bp
     1  ATCTTTTGCAGCTTTGCAGATACCCAGACTGAGCTGGAACTGGAATTTGTCTTCCTATTG
    61  ACTCTACTTCTTTAAAAGCGGCTGCCCATTACATTCCTCAGCTGTCCTTGCAGTTAGGTG
   121  TACATGTGACTGAGTGTTGGCCAGTGAGATGAAGTCTCCTCAAAGGAAGGCAGCATGTGT
   181  CCTTTTTCATCCCTTCATCTTGCTGCTGGGATTGTGGATATAACAGGAGCCCTGGCAGCT
   241  GTCTCCAGAGGATCAAAGCCACACCCAAAGAGTAAGGCAGATTAGAGACCAGAAAGACCT
   301  TGACTACTTCCCTACTTCCACTGCTTTTTCCTGCATTTAAGCCATTGTAAATCTGGGTGT
   361  GTTACATGAAGTGAAAATTAATTCTTTCTGCCCTTCAGTTCTTTATCCTGATACCATTTA
   421  ACACTGTCTGAATTAACTAGACTGCAATAATTCTTTCTTTTGAAAGCTTTTAAAGGATAA
   481  TGTGCAATTCACATTAAAATTGATTTTCCATTGTCAATTAGTTATACTCATTTTCCTGCC
   541  TTGATCTTTCATTAGATATTTTGTATCTGCTTGGAATATATTATCTTCTTTTTAACTGTG
   601  TAATTGGTAATTACTAAAACTCTGTAATCTCCAAAATATTGCTATCAAATTACACACCAT
   661  GTTTTCTATCATTCTCATAGATCTGCCTTATAAACATTTAAATAAAAAGTACTATTTA
```

FIGURE 4F

```
   1  GATCCAGCTTGGGTAGGCGGGGAAGCAGCTGGAGTGCGACCGCCGCGGCAGCCACCCTGC
  61  AACCGCCAGTCGGAGAGAGCTAAGGGCAAGTCCTGAGGTTGGGCCCAGGAGAAAGAAGGC
 121  AAGGAGACATTGTCCCAGGTAGGATGTGTCCCAGGATATTCTTGGTGATCTTGGAAGTGT
 181  CCGTATCATGGAATCAATCTCTATGATGGGAAGCCCTAAGAGCCTTAGTGAAACTTTTTT
 241  ACCTAATGGCATAAATGGTATCAAAGATGCAAGGAAGGTCACTGTAGGTGTGATTGGAAG
 301  TGGAGATTTTGCCAAATCCTTGACCATTCGACTTATTAGATGCGGCTATCATGTGGTCAT
 361  AGGAAGTAGAAATCCTAAGTTTGCTTCTGAATTTTTTCCTCATGTGGTAGATGTCACTCA
 421  TCATGAAGATGCTCTCACAAAAACAAATATAATATTTGTTGCTATACACAGAGAACATTA
 481  TACCTCCCTGTGGGACCTGAGACATCTGCTTGTGGGTAAAATCCTGATTGATGTGAGCAA
 541  TAACATGAGGATAAACCAGTACCCAGAATCCAATGCTGAATATTTGGCTTCATTATTCCC
 601  AGATTCTTTGATTGTCAAAGGATTTAATGTTGTCTCAGCTTGGGCACTTCAGTTAGGACC
 661  TAAGGATGCCAGCCGGCAGGTTTATATATGCAGCAACAATATTCAAGCGCGACAACAGGT
 721  TATTGAACTTGCCCGCCAGTTGAATTTCATTCCCATTGACTTGGGATCCTTATCATCAGC
 781  CAGAGAGATTGAAAATTTACCCCTACGACTCTTTACTCTCTGGAGAGGGCCAGTGGTGGT
 841  AGCTATAAGCTTGGCCACATTTTTTTTCCTTTATTCCTTTGTCAGAGATGTGATTCATCC
 901  ATATTCTAGAAACCAACAGAGTGACTTTTACAAAATTCCTATAGAGATTGTGAATAAAAC
 961  CTTACCTATAGTTGCCATTACTTTGCTCTCCCTAGTATACCTCGCAGGTCTTCTGGCAGC
1021  TGCTTATCAACTTTATTACGGCACCAAGTATAGGAGATTTCCACCTTGGTTGGAAACCTG
1081  GTTACAGTGTAGAAAACAGCTTGGATTACTAAGTTTTTTCTTCGCTATGGTCCATGTTGC
1141  CTACAGCCTCTGCTTACCGATGAGAAGGTCAGAGAGATATTTGTTTCTCAACATGGCTTA
1201  TCAGCAGGTTCATGCAAATATTGAAAACTCTTGGAATGAGGAAGAAGTTTGGAGAATTGA
1261  AATGTATATCTCCTTTGGCATAATGAGCCTTGGCTTACTTTCCCTCCTGGCAGTCACTTC
1321  TATCCCTTCAGTGAGCAATGCTTTAAACTGGAGAGAATTCAGTTTTATTCAGTCTACACT
1381  TGGATATGTCGCTCTGCTCATAAGTACTTTCCATGTTTTAATTTATGGATGGAAACGAGC
1441  TTTTGAGGAAGAGTACTACAGATTTTATACACCACCAAACTTTGTTCTTGCTCTTGTTTT
1501  GCCCTCAATTGTAATTCTGGATCTTTTGCAGCTTTGCAGATACCCAGACTGAGCTGGAAC
1561  TGGAATTTGTCTTCCTATTGACTCTACTTCTTTAAAAGCGGCTGCCCATTACATTCCTCA
1621  GCTGTCCTTGCAGTTAGGTGTACATGTGACTGAGTGTTGGCCAGTGAGATGAAGTCTCCT
1681  CAAAGGAAGGCAGCATGTGTCCTTTTTCATCCCTTCATCTTGCTGCTGGGATTGTGGATA
1741  TAACAGGAGCCCTGGCAGCTGTCTCCAGAGGATCAAAGCCACACCCAAAGAGTAAGGCAG
1801  ATTAGAGACCAGAAAGACCTTGACTACTTCCCTACTTCCACTGCTTTTTCCTGCATTTAA
1861  GCCATTGTAAATCTGGGTGTGTTACATGAAGTGAAAATTAATTCTTTCTGCCCTTCAGTT
1921  CTTTATCCTGATACCATTTAACACTGTCTGAATTAACTAGACTGCAATAATTCTTTCTTT
1981  TGAAAGCTTTTAAAGGATAATGTGCAATTCACATTAAAATTGATTTTCCATTGTCAATTA
2041  GTTATACTCATTTTCCTGCCTTGATCTTTCATTAGATATTTTGTATCTGCTTGGAATATA
2101  TTATCTTCTTTTTAACTGTGTAATTGGTAATTACTAAAACTCTGTAATCTCCAAAATATT
2161  GCTATCAAATTACACACCATGTTTTCTATCATTCTCATAGATCTGCCTTATAAACATTTA
2221  AATAAAAAGTACTATTTA
```

FIGURE 4G

```
   1 GATCCAGCTTGGGTAGGCGGGGAAGCAGCTGGAGTGCGACCGCCGCGGCAGCCACCCTGCA

  62 ACCGCCAGTCGGAGAGAGCTAAGGGCAAGTCCTGAGGTTGGGCCCAGGAGAAAGAAGGCA

 122 AGGAGACATTGTCCCAGGTAGGATGTGTCCCAGGATATTCTTGGTGATCTTGGAAGTGTC

           M  E  S  I  S  M  M  G  S  P  K  S  L  S  E  T  C  L
 182 CGTATCATGGAATCAATCTCTATGATGGGAAGCCCTAAGAGCCTTAGTGAAACTTGTTTA

  81  P  N  G  I  N  G  I  K  D  A  R  K  V  T  V  G  V  I  G  S
 242 CCTAATGGCATAAATGGTATCAAAGATGCAAGGAAGGTCACTGTAGGTGTGATTGGAAGT

 101  G  D  F  A  K  S  L  T  I  R  L  I  R  C  G  Y  H  V  V  I
 302 GGAGATTTTGCCAAATCCTTGACCATTCGACTTATTAGATGCGGCTATCATGTGGTCATA

 121  G  S  R  N  P  K  F  A  S  E  F  F  P  H  V  V  D  V  T  H
 362 GGAAGTAGAAATCCTAAGTTTGCTTCTGAATTTTTTCCTCATGTGGTAGATGTCACTCAT

 141  H  E  D  A  L  T  K  T  N  I  I  F  V  A  I  H  R  E  H  Y
 422 CATGAAGATGCTCTCACAAAAACAAATATAATATTTGTTGCTATACACAGAGAACATTAT

 161  T  S  L  W  D  L  R  H  L  L  V  G  K  I  L  I  D  V  S  N
 482 ACCTCCCTGTGGGACCTGAGACATCTGCTTGTGGGTAAAATCCTGATTGATGTGAGCAAT

 181  N  M  R  I  N  Q  Y  P  E  S  N  A  E  Y  L  A  S  L  F  P
 542 AACATGAGGATAAACCAGTACCCAGAATCCAATGCTGAATATTTGGCTTCATTATTCCCA

 201  D  S  L  I  V  K  G  F  N  V  V  S  A  W  A  L  Q  L  G  P
 602 GATTCTTTGATTGTCAAAGGATTTAATGTTGTCTCAGCTTGGGCACTTCAGTTAGGACCT

 221  K  D  A  S  R  Q  V  Y  I  C  S  N  N  I  Q  A  R  Q  Q  V
 662 AAGGATGCCAGCCGGCAGGTTTATATATGCAGCAACAATATTCAAGCGCGACAACAGGTT

 241  I  E  L  A  R  Q  L  N  F  I  P  I  D  L  G  S  L  S  S  A
 722 ATTGAACTTGCCCGCCAGTTGAATTTCATTCCCATTGACTTGGGATCCTTATCATCAGCC

 261  R  E  I  E  N  L  P  L  R  L  F  T  L  W  R  G  P  V  V  V
 782 AGAGAGATTGAAAATTTACCCCTACGACTCTTTACTCTCTGGAGAGGGCCAGTGGTGGTA

 281  A  I  S  L  A  T  F  F  F  L  Y  S  F  V  R  D  V  I  H  P
 842 GCTATAAGCTTGGCCACATTTTTTTTCCTTTATTCCTTTGTCAGAGATGTGATTCATCCA

 301  Y  A  R  N  Q  Q  S  D  F  Y  K  I  P  I  E  I  V  N  K  T
 902 TATGCTAGAAACCAACAGAGTGACTTTTACAAAATTCCTATAGAGATTGTGAATAAAACC

 321  L  P  I  V  A  I  T  L  L  S  L  V  Y  L  A  G  L  L  A  A
 962 TTACCTATAGTTGCCATTACTTTGCTCTCCCTAGTATACCTCGCAGGTCTTCTGGCAGCT

 341  A  Y  Q  L  Y  Y  G  T  K  Y  R  R  F  P  P  W  L  E  T  W
1022 GCTTATCAACTTTATTACGGCACCAAGTATAGGAGATTTCCACCTTGGTTGGAAACCTGG

 361  L  Q  C  R  K  Q  L  G  L  L  S  F  F  F  A  M  V  H  V  A
1082 TTACAGTGTAGAAAACAGCTTGGATTACTAAGTTTTTTCTTCGCTATGGTCCATGTTGCC

 381  Y  S  L  C  L  P  M  R  R  S  E  R  Y  L  F  L  N  M  A  Y
```

FIGURE 4H

```
1142  TACAGCCTCTGCTTACCGATGAGAAGGTCAGAGAGATATTTGTTTCTCAACATGGCTTAT

 401   Q  Q  V  H  A  N  I  E  N  S  W  N  E  E  E  V  W  R  I  E
1202  CAGCAGGTTCATGCAAATATTGAAAACTCTTGGAATGAGGAAGAAGTTTGGAGAATTGAA

 421   M  Y  I  S  F  G  I  M  S  L  G  L  L  S  L  L  A  V  T  S
1262  ATGTATATCTCCTTTGGCATAATGAGCCTTGGCTTACTTTCCCTCCTGGCAGTCACTTCT

 441   I  P  S  V  S  N  A  L  N  W  R  E  F  S  F  I  Q  S  T  L
1322  ATCCCTTCAGTGAGCAATGCTTTAAACTGGAGAGAATTCAGTTTTATTCAGTCTACACTT

 461   G  Y  V  A  L  L  I  S  T  F  H  V  L  I  Y  G  W  K  R  A
1382  GGATATGTCGCTCTGCTCATAAGTACTTTCCATGTTTTAATTTATGGATGGAAACGAGCT

 481   F  E  E  E  Y  Y  R  F  Y  T  P  P  N  F  V  L  A  L  V  L
1442  TTTGAGGAAGAGTACTACAGATTTTATACACCACCAAACTTTGTTCTTGCTCTTGTTTTG

 501   P  S  I  V  I  L  D  L  L  Q  L  C  R  Y  P  D  -
1502  CCCTCAATTGTAATTCTGGATCTTTTGCAGCTTTGCAGATACCCAGACTGAGCTGGAACT
```

FIGURE 4H

```
EXON_1  75bp
      1 GATCCAGCTTGGGTAGGCGGGGAAGCAGCTGGAGTGCGACCGCCGCGGCAGCCACCCTGC
     61 AACCGCCAGTCGGAG

EXON_2  79bp
      1 AGAGCTAAGGGCAAGTCCTGAGGTTGGGCCCAGGAGAAAGAAGGCAAGGAGACATTGTCC
     61 CAGGTAGGATGTGTCCCAG

EXON_3  525bp
      1 GATATTCTTGGTGATCTTGGAAGTGTCCGTATCATGGAATCAATCTCTATGATGGGAAGC
     61 CCTAAGAGCCTTAGTGAAACTTTTTTACCTAATGGCATAAATGGTATCAAAGATGCAAGG
    121 AAGGTCACTGTAGGTGTGATTGGAAGTGGAGATTTTGCCAAATCCTTGACCATTCGACTT
    181 ATTAGATGCGGCTATCATGTGGTCATAGGAAGTAGAAATCCTAAGTTTGCTTCTGAATTT
    241 TTTCCTCATGTGGTAGATGTCACTCATCATGAAGATGCTCTCACAAAAACAAATATAATA
    301 TTTGTTGCTATACACAGAGAACATTATACCTCCCTGTGGGACCTGAGACATCTGCTTGTG
    361 GGTAAAATCCTGATTGATGTGAGCAATAACATGAGGATAAACCAGTACCCAGAATCCAAT
    421 GCTGAATATTTGGCTTCATTATTCCCAGATTCTTTGATTGTCAAAGGATTTAATGTTGTC
    481 TCAGCTTGGGCACTTCAGTTAGGACCTAAGGATGCCAGCCGGCAG

EXON_4  528bp
      1 GTTTATATATGCAGCAACAATATTCAAGCGCGACAACAGGTTATTGAACTTGCCCGCCAG
     61 TTGAATTTCATTCCCATTGACTTGGGATCCTTATCATCAGCCAGAGAGATTGAAAATTTA
    121 CCCCTACGACTCTTTACTCTCTGGAGAGGGCCAGTGGTGGTAGCTATAAGCTTGGCCACA
    181 TTTTTTTTCCTTTATTCCTTTGTCAGAGATGTGATTCATCCATATGCTAGAAACCAACAG
    241 AGTGACTTTTACAAAATTCCTATAGAGATTGTGAATAAAACCTTACCTATAGTTGCCATT
    301 ACTTTGCTCTCCCTAGTATACCTCGCAGGTCTTCTGGCAGCTGCTTATCAACTTTATTAC
    361 GGCACCAAGTATAGGAGATTTCCACCTTGGTTGGAAACCTGGTTACAGTGTAGAAAACAG
    421 CTTGGATTACTAAGTTTTTTCTTCGCTATGGTCCATGTTGCCTACAGCCTCTGCTTACCG
    481 ATGAGAAGGTCAGAGAGATATTTGTTTCTCAACATGGCTTATCAGCAG

EXON_5  165bp
      1 GTTCATGCAAATATTGAAAACTCTTGGAATGAGGAAGAAGTTTGGAGAATTGAAATGTAT
     61 ATCTCCTTTGGCATAATGAGCCTTGGCTTACTTTCCCTCCTGGCAGTCACTTCTATCCCT
    121 TCAGTGAGCAATGCTTTAAACTGGAGAGAATTCAGTTTTATTCAG

EXON_7 and 3'UTR  718bp
      1 ATCTTTTGCAGCTTTGCAGATACCCAGACTGAGCTGGAACTGGAATTTGTCTTCCTATTG
     61 ACTCTACTTCTTTAAAAGCGGCTGCCCATTACATTCCTCAGCTGTCCTTGCAGTTAGGTG
    121 TACATGTGACTGAGTGTTGGCCAGTGAGATGAAGTCTCCTCAAAGGAAGGCAGCATGTGT
    181 CCTTTTTCATCCCTTCATCTTGCTGCTGGGATTGTGGATATAACAGGAGCCCTGGCAGCT
    241 GTCTCCAGAGGATCAAAGCCACACCCAAAGAGTAAGGCAGATTAGAGACCAGAAAGACCT
    301 TGACTACTTCCCTACTTCCACTGCTTTTTCCTGCATTTAAGCCATTGTAAATCTGGGTGT
    361 GTTACATGAAGTGAAAATTAATTCTTTCTGCCCTTCAGTTCTTTATCCTGATACCATTTA
    421 ACACTGTCTGAATTAACTAGACTGCAATAATTCTTTCTTTTGAAAGCTTTTAAAGGATAA
    481 TGTGCAATTCACATTAAAATTGATTTTCCATTGTCAATTAGTTATACTCATTTTCCTGCC
    541 TTGATCTTTCATTAGATATTTTGTATCTGCTTGGAATATATTATCTTCTTTTTAACTGTG
    601 TAATTGGTAATTACTAAAACTCTGTAATCTCCAAAATATTGCTATCAAATTACACACCAT
    661 GTTTTCTATCATTCTCATAGATCTGCCTTATAAACATTTAAATAAAAAGTACTATTTA
```

FIGURE 4I

```
   1  GGATCCAGCTTGGGTAGGCGGGGAAGCAGCTGGAGTGCGACCGCTACGGCAGCCACCCTG
  61  CAACCGCCAGTCGGAGAGCTAAGGGCAAGTCCTGAGGTTGGGCCCAGGAGAAAGAAGGCA
 121  AGGAGACATTGTCCCAGGATATTCTTGGTGATCTTGGAAGTGTCCGTATCATGGAATCAA
 181  TCTCTATGATGGGAAGCCCTAAGAGCCTTAGTGAAACTTGTTTACCTAATGGCATAAATG
 241  GTATCAAAGATGCAAGGAAGGTCACTGTAGGTGTGATTGGAAGTGGAGATTTTGCCAAAT
 301  CCTTGACCATTCGACTTATTAGATGCGGCTATCATGTGGTCATAGGAAGTAGAAATCCTA
 361  AGTTTGCTTCTGAATTTTTTCCTCATGTGGTAGATGTCACTCATCATGAAGATGCTCTCA
 421  CAAAAACAAATATAATATTTGTTGCTATACACAGAGAACATTATACCTCCCTGTGGGACC
 481  TGAGACATCTGCTTGTGGGTAAAATCCTGATTGATGTGAGCAATAACATGAGGATAAACC
 541  AGTACCCAGAATCCAATGCTGAATATTTGGCTTCATTATTCCCAGATTCTTTGATTGTCA
 601  AAGGATTTAATGTTGTCTCAGCTTGGGCACTTCAGTTAGGACCTAAGGATGCCAGCCGGC
 661  AGGTTTATATATGCAGCAACAATATTCAAGCGCGACAACAGGTTATTGAACTTGCCCGCC
 721  AGTTGAATTTCATTCCCATTGACTTGGGATCCTTATCATCAGCCAGAGAGATTGAAAATT
 781  TACCCCTACGACTCTTTACTTTCTGGAGAGGGCCAGTGGTGGTAGCTATAAGCTTGGCCA
 841  CATTTTTTTTCCTTTATTCCTTTGTCAGAGATGTGATTCATCCATATGCTAGAAACCAAC
 901  AGAGTGACTTTTACAAAATTCCTATAGAGATTGTGAATAAAACCTTACCTATAGTTGCCA
 961  TTACTTTGCTCTCCCTAGTATACCTTGCAGGTCTTCTGGCAGCTGCTTATCAACTTTATT
1021  ACGGCACCAAGTATAGGAGATTTCCACCTTGGTTGGAAACCTGGTTACAGTGTAGAAAAC
1081  AGCTTGGATTACTAAGTTTTTTCTTCGCTATGGTCCATGTTGCCTACAGCCTCTGCTTAC
1141  CGATGAGAAGGTCAGAGAGATATTTGTTTCTCAACATGGCTTATCAGCAGGTTCATGCAA
1201  ATATTGAAAACTCTTGGAATGAGGAAGAAGTTTGGAGAATTGAAATGTATATCTCCTTTG
1261  GCATAATGAGCCTTGGCTTACTTTCCCTCCTGGCAGTCACTTCTATCCCTTCGGTGAGCA
1321  ATGCTTTAAACTGGAGAGAATTCAGTTTTATTCAGATCTTTTGCAGCTTTGCAGATACCC
1381  AGACTGAGCTGGAACTGGAATTTGTCTTCCTATTGACTCTACTTCTTTAAAAGCGGCTGC
1441  CCATTACATTCCTCAGCTGTCCTTGCAGTTAGGTGTACATGTGACTGAGTGTTGGCCAGT
1501  GAGATGAAGTCTCCTCAAAGGAAGGCAGCATGTGTCCTTTTTCATCCCTTCATCTTGCTG
1561  CTGGGATTGTGGATATAACAGGAGCCCTGGCAGCTGCTCCAGAGGATCAAAGCCACACCC
1621  AAAGAGTAAGGCAGATTAGAGACCAGAAAGACCTTGACTACTTCCCTACTTCCACTGCTT
1681  TTTCCTGCATTTAAGCCATTGTAAATCTGGGTGTGTTACATGAAGTGAAAATTAATTCTT
1741  TCTGCCCTTCAGTTCTTTATCCTGATACCATTTAACACTGTCTGAATTAACTAGACTGCA
1801  ATAATTCTTTCTTTTGAAAGCTTTTAAAGGATAATGTGCAATTCACATTAAAATTGATTT
1861  TCCATTGTCAATTAGTTATACTCATTTTCCTGCCTTGATCTTTCATTAGATATTTTGTAT
1921  CTGCTTGGAATATATTATCTTCTTTTTAACTGTGTAATTGGTAATTACTAAAACTCTGTA
1981  ATCTCCAAAATATTGCTATCAAATTACACACCATGTTTTCTATCATTCTCATAGATCTGC
2041  CTTATAAACATTTAAATAAAAAGTACTATTTACCAAAAAAAAAAAAAAAAAAAAAAAAA
2101  AA
```

FIGURE 4J

```
1143  ATGAGAAGGTCAGAGAGATATTTGTTTCTCAACATGGCTTATCAGCAGGTTCATGCAAAT

 345   I  E  N  S  W  N  E  E  E  V  W  R  I  E  M  Y  I  S  F  G
1203  ATTGAAAACTCTTGGAATGAGGAAGAAGTTTGGAGAATTGAAATGTATATCTCCTTTGGC

 365   I  M  S  L  G  L  L  S  L  L  A  V  T  S  I  P  S  V  S  N
1263  ATAATGAGCCTTGGCTTACTTTCCCTCCTGGCAGTCACTTCTATCCCTTCGGTGAGCAAT

 385   A  L  N  W  R  E  F  S  F  I  Q  I  F  C  S  F  A  D  T  Q
1323  GCTTTAAACTGGAGAGAATTCAGTTTTATTCAGATCTTTTGCAGCTTTGCAGATACCCAG

 405   T  E  L  E  L  E  F  V  F  L  L  T  L  L  L  *
1383  ACTGAGCTGGAACTGGAATTTGTCTTCCTATTGACTCTACTTCTTTAAAAGCGGCTGCCC
1443  ATTACATTCCTCAGCTGTCCTTGCAGTTAGGTGTACATGTGACTGAGTGTTGGCCAGTGA
1503  GATGAAGTCTCCTCAAAGGAAGGCAGCATGTGTCCTTTTTCATCCCTTCATCTTGCTGCT
1563  GGGATTGTGGATATAACAGGAGCCCTGGCAGCTGCTCCAGAGGATCAAAGCCACACCCAA
1623  AGAGTAAGGCAGATTAGAGACCAGAAAGACCTTGACTACTTCCCTACTTCCACTGCTTTT
1683  TCCTGCATTTAAGCCATTGTAAATCTGGGTGTGTTACATGAAGTGAAAATTAATTCTTTC
1743  TGCCCTTCAGTTCTTTATCCTGATACCATTTAACACTGTCTGAATTAACTAGACTGCAAT
1803  AATTCTTTCTTTTGAAAGCTTTTAAAGGATAATGTGCAATTCACATTAAAATTGATTTTC
1863  CATTGTCAATTAGTTATACTCATTTTCCTGCCTTGATCTTTCATTAGATATTTTGTATCT
1923  GCTTGGAATATATTATCTTCTTTTTAACTGTGTAATTGGTAATTACTAAAACTCTGTAAT
1983  CTCCAAAATATTGCTATCAAATTACACACCATGTTTTCTATCATTCTCATAGATCTGCCT
2043  TATAAACATTTAAATAAAAAGTACTATTTACCAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 4K

```
1083  TTTGTACTCTTGGGAATCACTTCTTTGCCATCTGTTAGCAATGCAGTCAACTGGAGAGAG

 379   F  R  F  V  Q  S  K  L  G  Y  L  T  L  I  L  C  T  A  H  T
1143  TTCCGATTTGTCCAGTCCAAACTGGGTTATTTGACCCTGATCTTGTGTACAGCCCACACC

 399   L  V  Y  G  G  K  R  F  L  S  P  S  N  L  R  W  Y  L  P  A
1203  CTGGTGTACGGTGGGAAGAGATTCCTCAGCCCTTCAAATCTCAGATGGTATCTTCCTGCA

 429   A  Y  V  L  G  L  I  I  P  C  T  V  L  V  I  K  F  V  L  I
1263  GCCTACGTGTTAGGGCTTATCATTCCTTGCACTGTGCTGGTGATCAAGTTTGTCCTAATC

 439   M  P  C  V  D  N  T  L  T  R  I  R  Q  G  W  E  R  N  S  K
1323  ATGCCATGTGTAGACAACACCCTTACAAGGATCCGCCAGGGCTGGGAAAGGAACTCAAAA

 459   H  -
1383  CACTAGCTCGAGGT
```

FIGURE 4L

```
1ACCCTTCGCCGCGGACCTTCAGCTGCCGCGGTCGCTCCGAGCGGCGGGCCGCAGAGGTTC

63  AAGCGATTCTCCTGCTTCAGCCTCCGGAGTAGCTGGGATTACAGGCACGTGCCAACACAC

 1                M  P  E  E  M  D  K  P  L  I  S  L  H  L  V  D
123 CCAGCCACCAAAATGCCAGAAGAGATGGACAAGCCACTGATCAGCCTCCACCTGGTGGAC

17   S  D  S  S  L  A  K  V  P  D  E  A  P  K  V  G  I  L  G  S
183 AGCGATAGTAGCCTTGCCAAGGTCCCCGATGAGGCCCCCAAAGTGGGCATCCTGGGTAGC

37   G  D  F  A  R  S  L  A  T  R  L  V  G  S  G  F  K  V  V  V
243 GGGGACTTTGCCCGCTCCCTGGCCACACGCCTGGTGGGCTCTGGCTTCAAAGTGGTGGTG

57   G  S  R  N  P  K  R  T  A  R  L  Y  P  S  A  A  Q  V  T  F
303 GGGAGCCGCAACCCCAAACGCACAGCCAGGCTGTATCCCTCAGCGGCCCAAGTGACTTTC

77   Q  E  E  A  V  S  S  P  E  V  I  F  V  A  V  F  R  E  H  Y
363 CAAGAGGAGGCAGTGAGCTCCCCGGAGGTCATCTTTGTGGCTGTGTTCCGGGAGCACTAC

97   S  S  L  C  S  L  S  D  Q  L  A  G  K  I  L  V  D  V  S  N
423 TCTTCACTGTGCAGTCTCAGTGACCAGCTGGCGGGCAAGATCCTGGTGGATGTGAGCAAC

117  P  T  E  Q  E  H  L  Q  H  R  E  S  N  A  E  Y  L  A  S  L
483 CCTACAGAGCAAGAGCACCTTCAGCATCGTGAGTCCAATGCTGAGTACCTGGCCTCCCTC

137  F  P  T  C  T  V  V  K  A  F  N  V  I  S  A  W  T  L  Q  A
543 TTCCCCACTTGCACAGTGGTCAAGGCCTTCAATGTCATCTCTGCCTGGACCCTGCAGGCT

157  G  P  R  D  G  N  R  Q  V  P  I  C  G  D  Q  P  E  A  K  R
603 GGCCCAAGGGATGGTAACAGGCAGGTGCCCATCTGCGGTGACCAGCCAGAAGCCAAGCGT

177  A  V  S  E  M  A  L  A  M  G  F  M  P  V  D  M  G  S  L  A
663 GCTGTCTCGGAGATGGCGCTCGCCATGGGCTTCATGCCCGTGGACATGGGATCCCTGGCG

197  S  A  W  E  V  E  A  M  P  L  R  L  L  P  A  W  K  V  P  T
723 TCAGCCTGGGAGGTGGAGGCCATGCCCCTGCGCCTCCTCCCGGCCTGGAAGGTGCCCACC

217  L  L  A  L  G  L  F  V  C  F  Y  A  Y  N  F  V  R  D  V  L
783 CTGCTGGCCCTGGGGCTCTTCGTCTGCTTCTATGCCTACAACTTCGTCCGGGACGTTCTG

237  Q  P  Y  V  Q  E  S  Q  N  K  F  F  K  L  P  V  S  V  V  N
843 CAGCCCTATGTGCAGGAAAGCCAGAACAAGTTCTTCAAGCTGCCCGTGTCCGTGGTCAAC

257  T  T  L  P  C  V  A  Y  V  L  L  S  L  V  Y  L  P  G  V  L
903 ACCACACTGCCGTGCGTGGCCTACGTGCTGCTGTCACTCGTGTACTTGCCCGGCGTGCTG

277  A  A  A  L  Q  L  R  R  G  T  K  Y  Q  R  F  P  D  W  L  D
963 GCGGCTGCCCTGCAGCTGCGGCGCGGCACCAAGTACCAGCGCTTCCCCGACTGGCTGGAC

297  H  W  L  Q  H  R  K  Q  I  G  L  L  S  F  F  C  A  A  L  H
023 CACTGGCTACAGCACCGCAAGCAGATCGGGCTGCTCAGCTTCTTCTGCGCCGCCCTGCAC

317  A  L  Y  S  F  C  L  P  L  R  R  A  H  R  Y  D  L  V  N  L
.083 GCCCTCTACAGCTTCTGCTTGCCGCTGCGCCGCGCCCACCGCTACGACCTGGTCAACCTG

337  A  V  K  Q  V  L  A  N  K  S  H  L  W  V  E  E  E  V  W  R
.143 GCAGTCAAGCAGGTCTTGGCCAACAAGAGCCACCTCTGGGTGGAGGAGGAGGTCTGGCGG

357  M  E  I  Y  L  S  L  G  V  L  A  L  G  T  L  S  L  L  A  V
.203 ATGGAGATCTACCTCTCCCTGGGAGTGCTGGCCCTCGGCACGTTGTCCCTGCTGGCCGTG
```

FIGURE 4M

```
 377   T  S  L  P  S  I  A  N  S  L  N  W  R  E  F  S  F  V  Q  S
1263   ACCTCACTGCCGTCCATTGCAAACTCGCTCAACTGGAGGGAGTTCAGCTTCGTTCAGTCC

 397   S  L  G  F  V  A  L  V  L  S  T  L  H  T  L  T  Y  G  W  T
1323   TCACTGGGCTTTGTGGCCCTCGTGCTGAGCACACTGCACACGCTCACCTACGGCTGGACC

 417   R  A  F  E  E  S  R  Y  K  F  Y  L  P  P  T  F  T  L  T  L
1383   CGCGCCTTCGAGGAGAGCCGCTACAAGTTCTACCTGCCTCCCACCTTCACGCTCACGCTG

 437   L  V  P  C  V  V  I  L  A  K  A  L  F  L  L  P  C  I  S  R
1443   CTGGTGCCCTGCGTCGTCATCCTGGCCAAAGCCCTGTTTCTCCTGCCCTGCATCAGCCGC

 457   R  L  A  R  I  R  R  G  W  E  R  E  S  T  I  K  F  T  L  P
1503   AGACTCGCCAGGATCCGGAGAGGCTGGGAGAGGGAGAGCACCATCAAGTTCACGCTGCCC

 477   T  D  H  A  L  A  E  K  T  S  H  V  -
1563   ACAGACCACGCCCTGGCCGAGAAGACGAGCCACGTATGAGGTGCCTGCCCTGGGCTCTGG

1623   ACCCCGGGCACACGAGGGACGGTGCCCTGAGCCCGTTAGGTTTTCTTTTCTTGGTGGTGC
1683   AAAGTGGTATAACTGTGTGCAAATAGGAGGTTTGAGGTCCAAATTCCTGGGACTCAAATG
1743   TATGCAGTACTATTCAGAATGATATACACACATATGTGTATATGTATTTACATATATTCC
1803   ACATATATAACAGGATTTGCAATTATACATAGCTAGCTAAAAAGTTGGGTCTCTGAGATT
1863   TCAACTTGTAGATTTAAAAACAAGTGCCGTACGTTAAGAGAAGAGCAGATCATGCTATTG
1923   TGACATTTGCAGAGATATACACACACTTTTTGTACAGAAGAGGCTTGTGCTGTGGTGGGT
1983   TCGATTTATCCCTGCCCACCCCATCCCCACAACTTCCCTTTTGCTACTTCCCCAAGGCTC
2043   TTGCAGAGCTAGGGCTCTGAAGGGGAGGGAAGGCAACGGCTCTGCCCAGAGCCATCCCTG
2103   GAGCATGTGAGCAGCGGCTGGTCTCTTCCCTCCACCTGGGGCAGCAGCAGGAGGCCTGGG
2163   GGGGAGGAAAATCAGGCAGTCGGCCTGGAGTCTGTGCCTGGTCCTTTGCCCGGTGGTGGG
2223   AGGATGGAGGGATTGGGCTGAAGCTGCTCCACCTCATCCTTGCTGAGTGGGGGAGACATT
2283   TTCCCTGAAAGTCAGAAGTCACCATAGAGCCTGCAAATGGATCCTCCTGTGAGAGTGACG
2343   TCACCTCCTTTCCAGAGCCATTAGTGAGCCTGGCTTGGGAACAAGTGTAATTTCCTTCCC
2403   TCCTTTAACCTGGCGATGAGCGTCCTTTAAACCACTGTGCCTTCTCACCCTTTCCATCTT
2463   CAGTTTGAACGACTCCCAGGAAGGCCTAGAGCAGACCCTTTAGAAATCAGCCCAAGGGGG
2523   AGAGCAAGAGAAAACACTCTAGGGAGTAAAGCTCCCCGGGCGTCAGAGTTGAGCCCTGCC
2583   TGGGCTGAAGGACTGTCTTCACGAAGTCAGTCCTGAGGAAAAATATTGGGGACTCCAAAT
2643   GTCCTCTGGCAGAGGACCCAGAAAACCACACTGGCTCCAACTTCCTCCTCATGGGGCATT
2703   ACACTTCAAAACAGTGGGGAGCAACTTTTCCACCAAAGCTACAAACCTAAAATGCTGCTG
2763   CCCCAAAGCACAAGAGGGAAGAGCACCGCCGGGGCCACAGGACGTCTGTCCTCCAGTCAC
2823   AGGCCATCCTTGCTGCTCCCTACTGACTCTAGCTTACTTCCCCTGTGAAGAAACAGGTGT
2883   TCTCGGCTGAGCCCCCAACCCTCTGCAGAACCAGGTTGATCTGCCACAGAAAAAGCATCT
2943   TTGAAGACAAAGAGGGTGAGGTCTTCATGAGTCTCCTGGGCCCAAAGCCATCTTCTGATG
3003   GAAGGAAGAGAGTAGGGCCAGTGAAGGCTGCCCAGAGAGAATGTCACAGATGAGGCTGCC
3063   CCTGCCCCCTCCCCGCCAGGGAGGTTTCATGAGCTCATGTCTATGCAGCACATAAGGGTT
3123   CTTCAGTGAAAAGCAGGAGAAGAGCCCACTGCAAGGATAGCTCATTAGGCACATGACCGA
3183   TGCAGGGAAGGCCATGCCGGGGAAGCTCTTCCTGCAGGTATTTTCCATCTGCTGTGCCAA
3243   GGCTGAGCGGCAGAAACTTGTCTCATAAATTGGCACTGATGGAGCATCAGCTGTGGCCCA
3303   CAGAGAGCCTTGCTGAGAAGGGGGCAGGTAAAGCAGAGATTTTAGCATTGCCTTGGCATA
3363   ACAAGGGCCCATCGATTCCCTACTAATGAGAGGCAGGGAGAGCATGGGCAATGGAGACCC
3423   ACCAATGATCCCCAACCCCGGTGGGTACTGGCTGCCTGCCCTGGGCCAGGGAATGGCTCC
3483   TTATACCAAAGATGCTGGCACATAGCAGAACCCAGTGCACGTCCTCCCCTTCCCACCCAC
3543   CTCTGGCTGAAGGTGCTCAAGAGGGAAGCAATTATAAGGTGGGTGGCAGGAGGGAACAGG
3603   TGCCACCTGCTGGACAATCACACGAAAGGCAGGCGGGCTGTGTACTGGGCCCTGACTGTG
3663   CGTCCACTGCTGTCTTCCCTACCTCACCAGGCTACTGGCAGCAGCATCCCGAGAGCACAT
3723   CATCTCCACAGCCTGGTAAATTCCATGTGCCTCTGGGTACAAAAGTGCCTCAACGACATG
3783   CTCTGGAAATCCCAAATGCCACAGTCTGAGGTTGATATCTAAAATCTATGCCTTCAAAAG
3843   AGTCTCTGTTTTTTTTTTTTAACCTGGTAGACGGTATAAAAGCAGTGCAAATAAACACCT
3903   AACCTTCTGC
```

FIGURE 4M

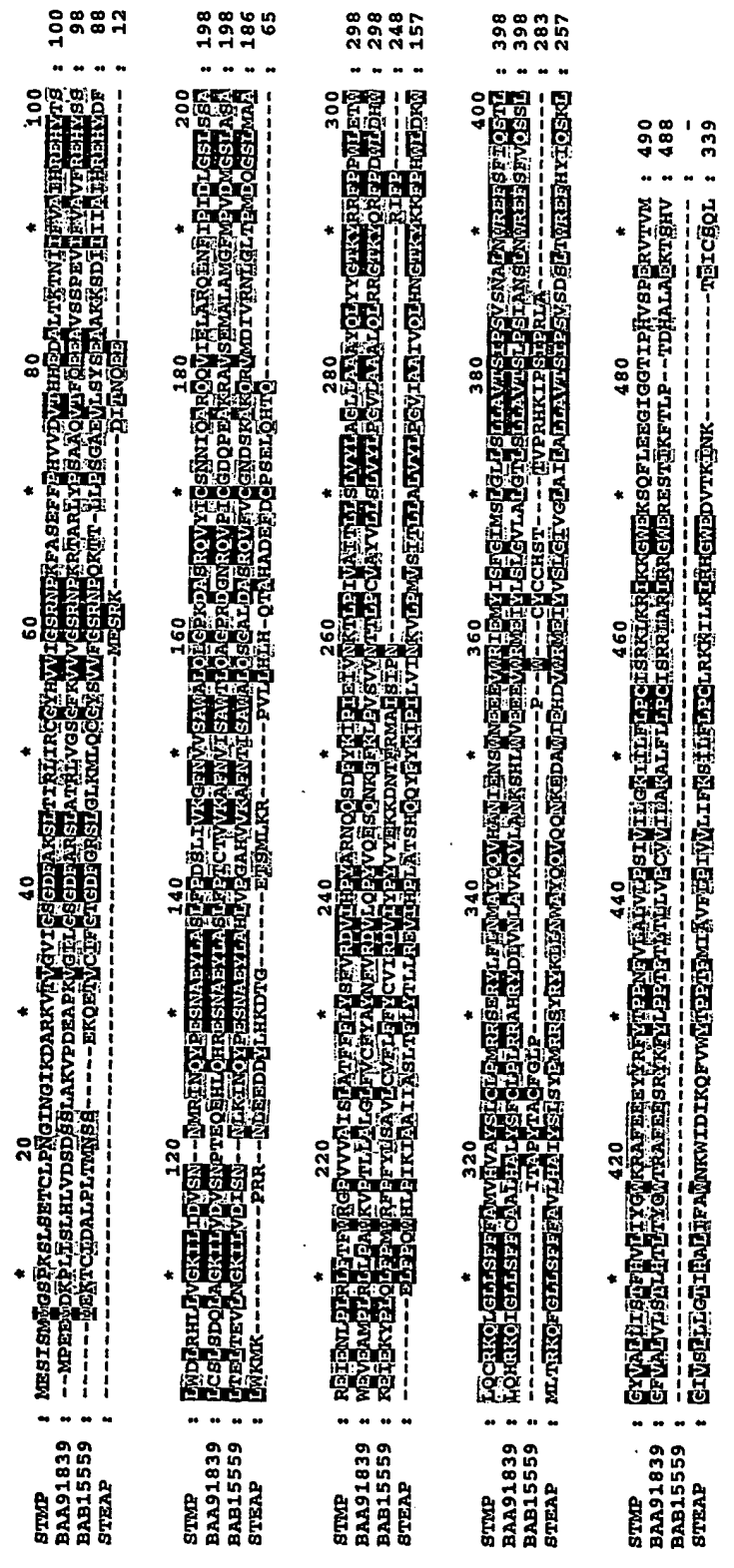
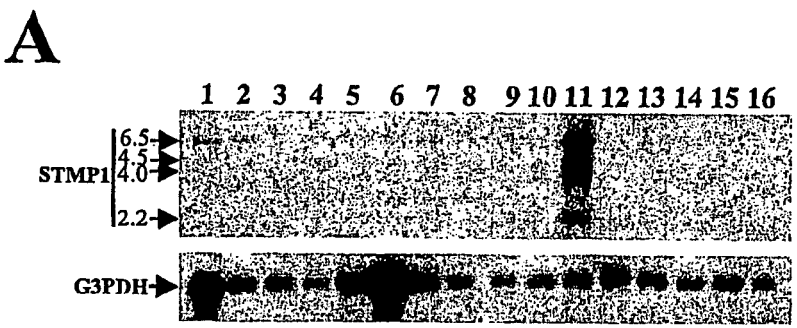


Figure 5


A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
STMP1
6.5
4.5
4.0
2.2
G3PDH

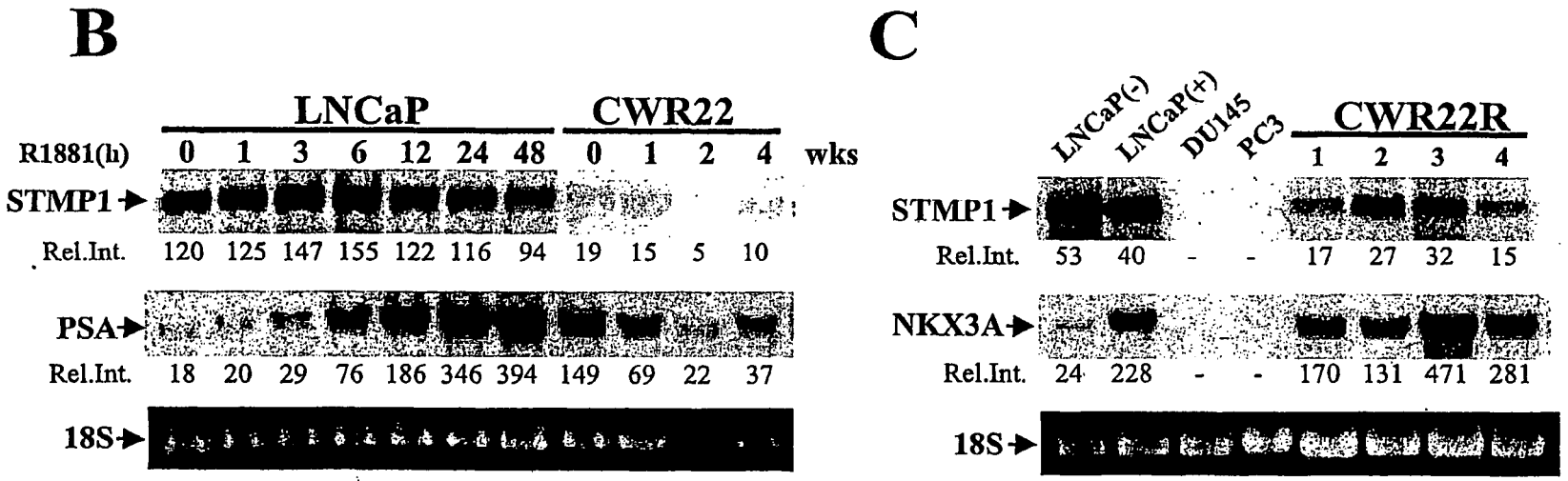

B
LNCaP
CWR22
R1881(h) 0 1 3 6 12 24 48 0 1 2 4 wks
STMP1
Rel.Int. 120 125 147 155 122 116 94 19 15 5 10
PSA
Rel.Int. 18 20 29 76 186 346 394 149 69 22 37
18S
C
LNCaP(-)
LNCaP(+)
DU145
PC3
CWR22R
1 2 3 4
STMP1
Rel.Int. 53 40 - - 17 27 32 15
NKX3A
Rel.Int. 24 228 - - 170 131 471 281
18S

Figure 6

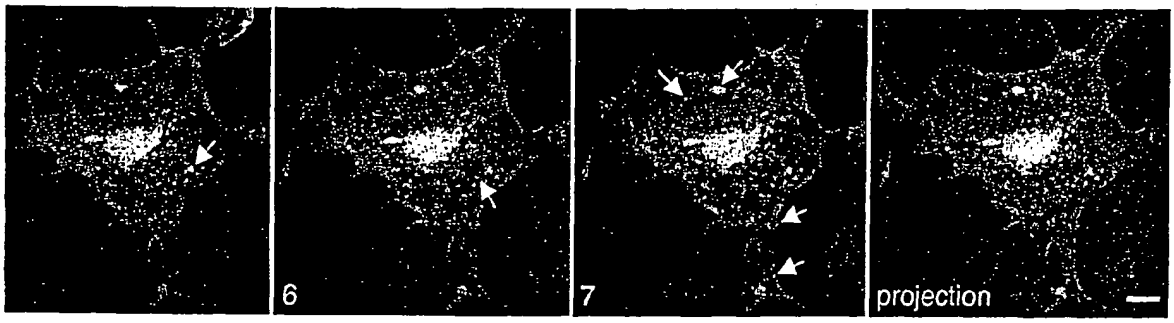


Figure 7A

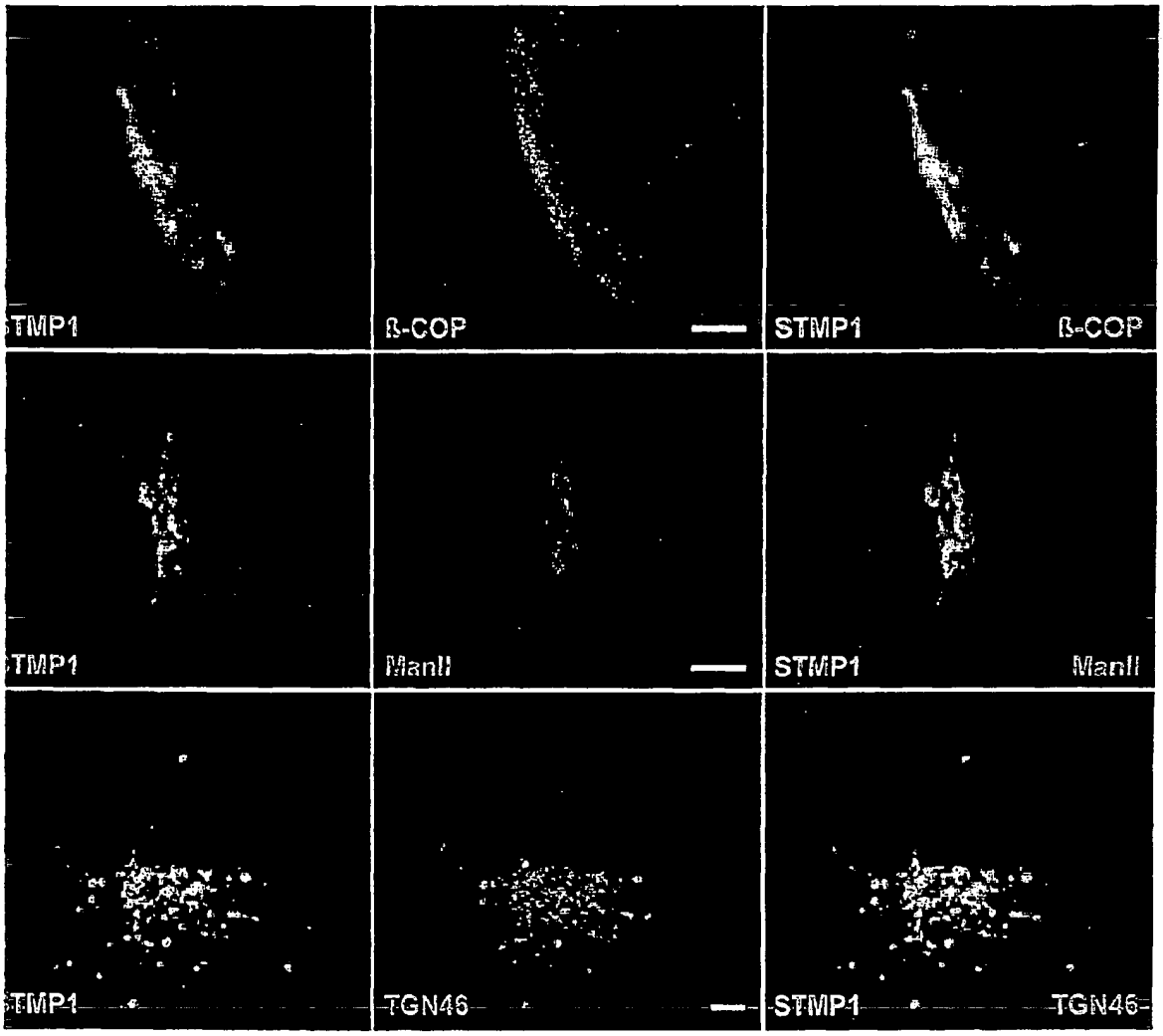

TMP1
ß-COP
STMP1
ß-COP
TMP1
ManII
STMP1
ManII
TMP1
TGN46
STMP1
TGN46

Figure 7B

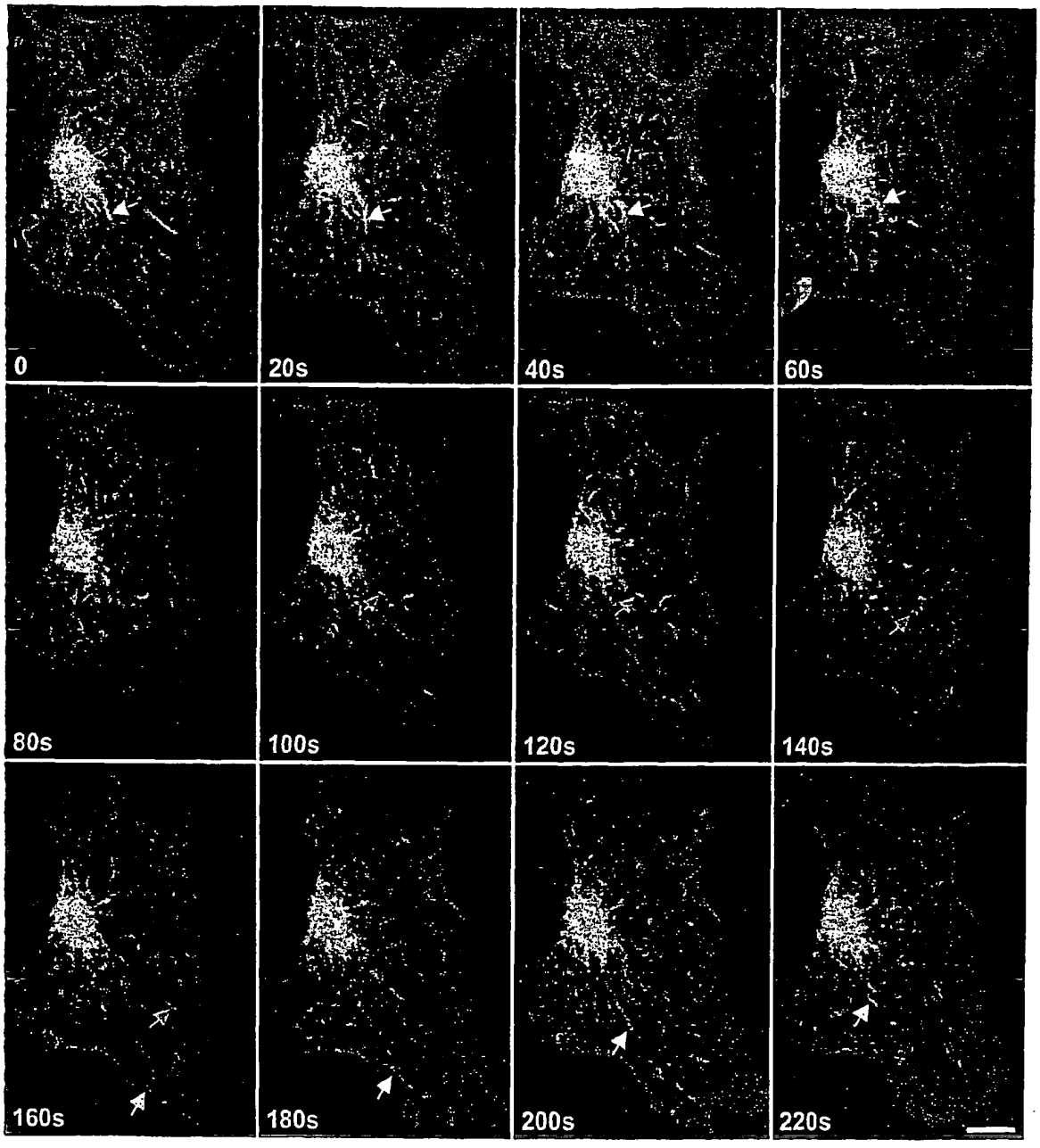


Figure 8

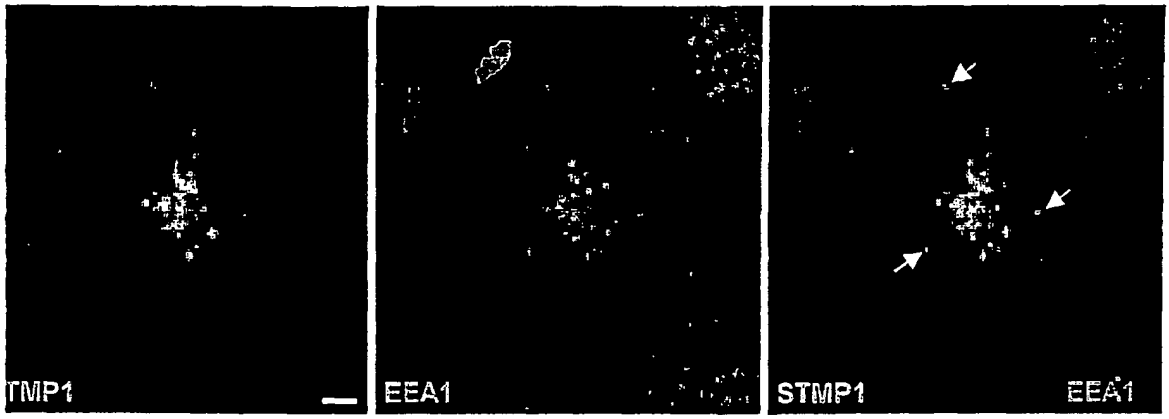


Figure 9

Two mRNA species are transcribed from the SSH9 gene

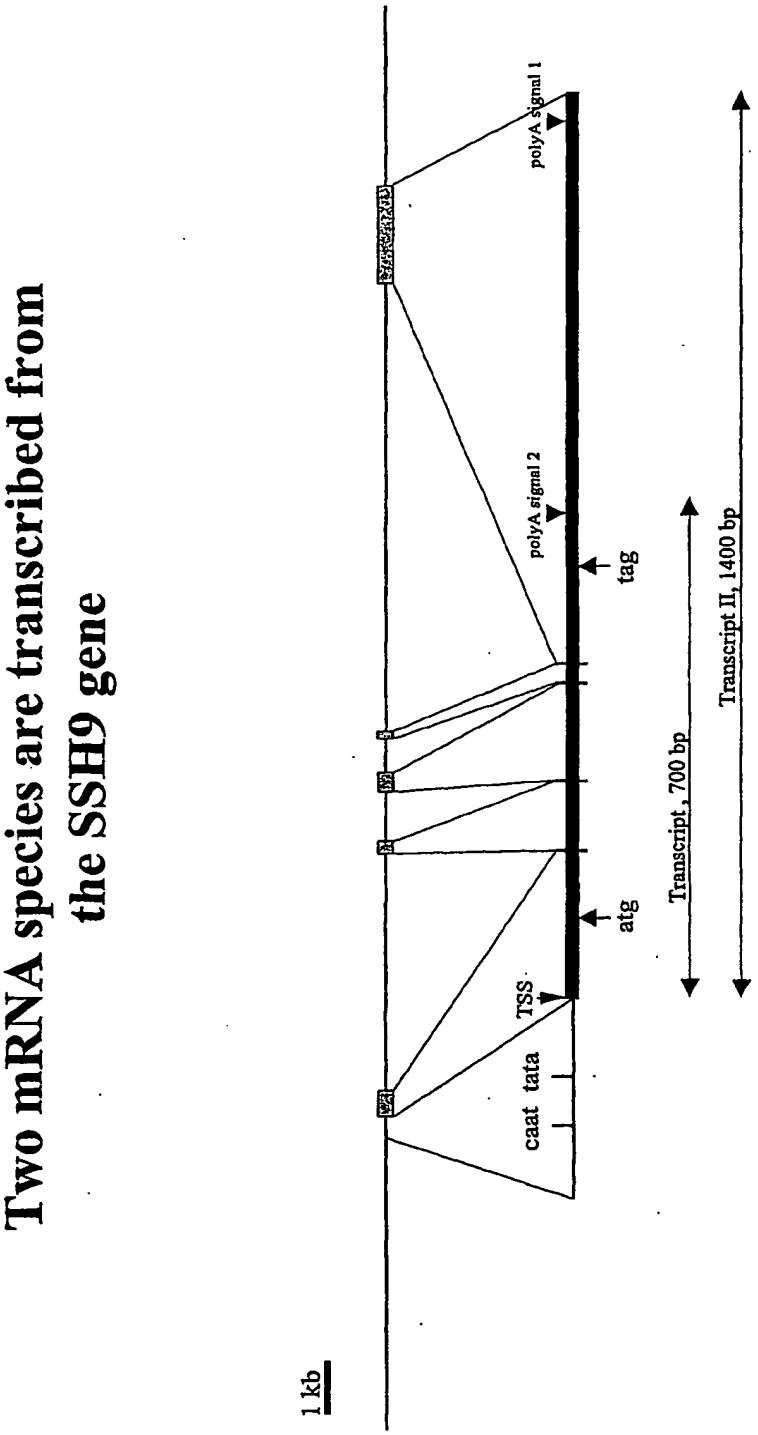


Figure 10

```
   1 AGCGGCGGCTCCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGG
   1                                                          M  T
  62 TCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCCATGACC

   3   T  T  T  T  F  K  G  V  D  P  N  S  R  N  S  S  R  V  L  R
 122 ACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTTGCGG

  23   P  P  G  G  G  S  N  F  S  L  G  F  D  E  P  T  E  Q  P  V
 182 CCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTGTG

  43   R  K  N  K  M  A  S  N  I  F  G  T  P  E  E  N  Q  A  S  W
 242 AGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

  63   A  K  S  A  G  A  K  S  S  G  G  R  E  D  L  E  S  S  G  L
 302 GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTG

  83   Q  R  R  N  S  S  E  A  S  S  G  D  F  L  D  L  K  G  E  G
 362 CAGAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGT

 103   D  I  H  E  N  V  D  T  D  L  P  G  S  L  G  Q  S  E  E  K
 422 GATATTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAG

 123   P  V  P  A  A  P  V  P  S  P  V  A  P  A  P  V  P  S  R  R
 482 CCCGTGCCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGA

 143   N  P  P  G  G  K  S  S  L  V  L  G  *
 542 AATCCCCCTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAGCTCTGACTGTCCTGAACGCTG

 602 TCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACTTGAGCCTGACTGTACATCT
 662 CTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTGTCTTTTTTTTTTTTCTT
 722 TTGAAGAACAGGTTTCTCTCTGTCCTTGACTCTTGGGTCTGTGGGCCATGGCATGAGTGT
 782 TTTCTAGTAGTAGATTGGAGGGAAAGCTTTGTGACACTTAGTACTGTGTTTTTAAGAAGA
 842 AATAATTTGGTTCCAGATGTGTTAGAGGATCTTTTGTACTGAGGTTTTTAACACTTTACT
 902 TGGGTTTACCAAGCCTCAACTGGACAGACCATAAACAGTCCACAGGCACCGTTCCTGCCA
 962 GGCCCCAACCCACAGGGAGTCTCTCCGCAGAGCCTTCTTGGTGTTGCCCTAACTTGCCAG
1022 TGGCCTTTGCTCAGAGCCTCCTCCTGTGACATGTGAACAATGAAGAGGCCTGCGCCTCCT
1082 GCCTTGCCGCCTGCAAAGCAAAGAAACTGCCTTTTATTTTTTAACCTTAAAAAGTAGCCA
1142 GATAGTAACAAGACTGGCTGGCTGATGAGCAAAGCCTTTGCTCTCACGCAGAGGAAGGCT
1202 TGGATGTACAATGAAACTGCCTGGAACTAAAAGCAGTGAAGCAAGGGAGGCAATCACACT
1262 GAAGCGGGTCTTCCTCCAGGAACGGGGTCCCACAGGCGTGTTGTTTTAAATAACCTGATG
1322 CTGTGTGCATGATGCTGGTGCTTGACCATGAAAGGAAAGTCTCATCCTTAAAATGTGTTG
1382 TACTTCACAATCCTGGACTGTTGCTTCAAGTAAACAATATCCACATTCTAAAAAAAAAAA
1442 AAAAAAAAAAAAAAAAAAAA
```

FIGURE 11A

FIGURE 11B

Predicted promoter

tgaaaaccctataaaggcgtcgatcggccggacaggcggcAgcggcggct

SSH9 EXON-Intron boundaries;

EXON1 CATGACCACAACCaccaccttcaaggga... INT1 ...tgccattatttgcagAGTTTTGCGGCCT

EXON2 AAATCAAGCTTCTtgggccaagtcagca... INT2 ...tattttgatttttagGTGCCAAGTCTAG

EXON3 CTTAGATCTGAAGgtcagtgtgacagca... INT4 ...ttttttctttttctagGGAGAAG

EXON4 GTGATATTCATGgtaagtacttctgaa... INT5 ...tccctgttttcatagAAAATGTGGACAC

FIGURE 11C

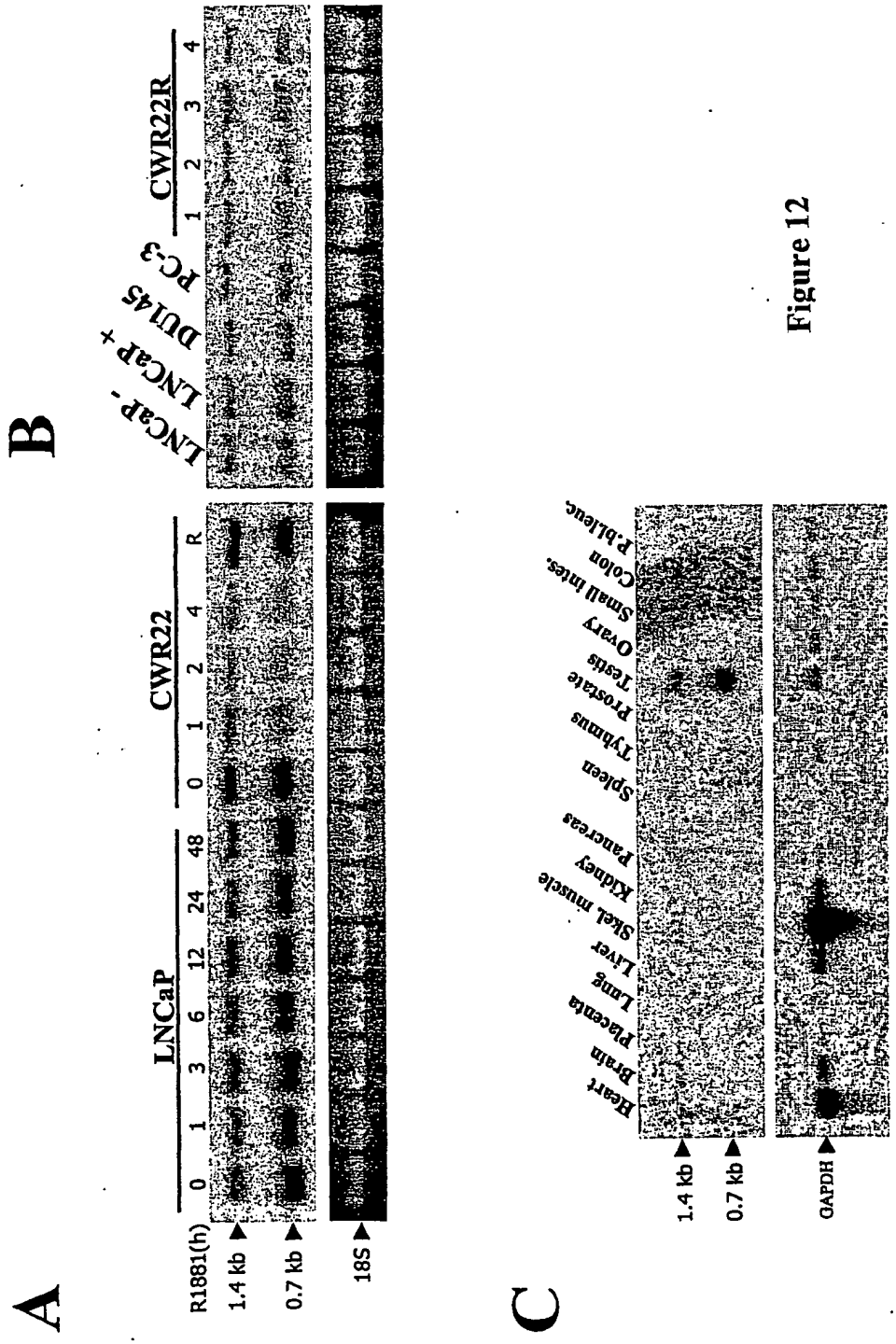

A
LNCaP
CWR22
R1881(h)
0 1 3 6 12 24 48 0 1 2 4 R
1.4 kb
0.7 kb
18S
B
LNCaP -
LNCaP +
DU145
PC-3
CWR22R
1 2 3 4
C
Heart
Brain
Placenta
Lung
Liver
Skel. muscle
Kidney
Pancreas
Spleen
Thymus
Prostate
Testis
Ovary
Small Intes.
Colon
P.bl.leuc.
1.4 kb
0.7 kb
GAPDH

Figure 12

PSL22 gene structure

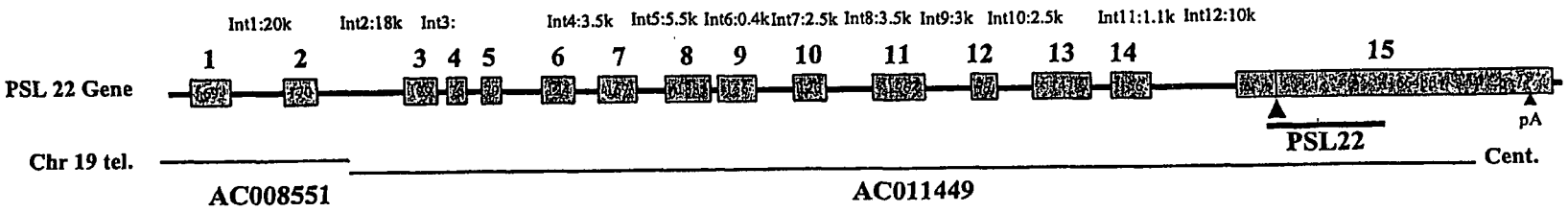


Figure 13

```
   1  ATGACCGACGCGCTGTTGCCCGCGGCCCCCCAGCCGCTGGAGAAGGAGAACGACGGCTAC
  61  TTTCGGAAGGGCTGTAATCCCCTTGCACAAACCGGCCGGAGTAAATTGCAGAATCAAAGA
 121  GCTGCTTTGAATCAGCAGATCCTGAAAGCCGTGCGGATGAGGACCGGAGCGGAAAACCTT
 181  CTGAAAGTGGCCACAAACTCAAAGGTGCGGGAGCAAGTGCGGCTGGAGCTGAGCTTCGTC
 241  AACTCAGACCTGCAGATGCTCAAGGAAGAGCTGGAGGGGCTGAACATCTCGGTGGGCGTC
 301  TATCAGAACACAGAGGAGGCATTTACGATTCCCCTGATTCCTCTTGGCCTGAAGGAAACG
 361  AAAGACGTCGACTTTGCAGTCGTCCTCAAGGATTTTATCCTGGAACATTACAGTGAAGAT
 421  GGCTATTTATATGAAGATGAAATTGCAGATCTTATGGATCTGAGACAAGCTTGTCGGACG
 481  CCTAGCCGGGATGAGGCCGGGGTGGAACTGCTGATGACATACTTCATCCAGCTGGGCTTT
 541  GTCGAGAGTCGATTCTTCCCGCCCACACGGCAGATGGGACTCCTGTTCACCTGGTATGAC
 601  TCTCTCACTGGGGTTCCGGTCAGCCAGCAGAACCTGCTGCTGGAGAAGGCCAGTGTCCTG
 661  TTCAACACTGGGGCCCTCTACACCCAGATTGGGACCCGGTGCGATCGGCAGACGCAGGCT
 721  GGGCTGGAGAGTGCCATAGATGCCTTTCAGAGAGCCGCAGGGGTTTTAAATTACCTGAAA
 781  GACACATTTACCCATACTCCAAGTTACGACATGAGCCCTGCCATGCTCAGCGTGCTCGTC
 841  AAAATGATGCTTGCACAAGCCCAAGAAAGCGTGTTTGAGAAAATCAGCCTTCCTGGGATC
 901  CGGAATGAATTCTTCATGCTGGTGAAGGTGGCTCAGGAGGCTGCTAAGGTGGGAGAGGTC
 961  TACCAACAGCTACACGCAGCCATGAGCCAGGCGCCGGTGAAAGAGAACATCCCCTACTCC
1021  TGGGCCAGCTTAGCCTGCGTGAAGGCCCACCACTACGCGGCCCTGGCCCACTACTTCACT
1081  GCCATCCTCCTCATCGACCACCAGGTGAAGCCAGGCACGGATCTGGACCACCAGGAGAAG
1141  TGCCTGTCCCAGCTCTACGACCACATGCCAGAGGGGCTGACACCCTTGGCCACACTGAAG
1201  AATGATCAGCAGCGCCGACAGCTGGGGAAGTCCCACTTGCGCAGAGCCATGGCTCATCAC
1261  GAGGAGTCGGTGCGGGAGGCGAGCCTCTGCAAGAAGCTGCGGAGCATTGAGGTGCTACAG
1321  AAGGTGCTGTGTGCCGCACAGGAACGCTCCCGGCTCACGTACGCCCAGCACCAGGAGGAG
1381  GATGACCTGCTGAACCTGATCGACGCCCCCAGTGTTGTTGCTAAAACTGAGCAAGAGGTT
1441  GACATTATATTGCCCCAGTTCTCCAAGCTGACAGTCACGGACTTCTTCCAGAAGCTGGGC
1501  CCCTTATCTGTGTTTTCGGCTAACAAGCGGTGGACGCCTCCTCGAAGCATCCGCTTCACT
1561  GCAGAAGAAGGGGACTTGGGGTTCACCTTGAGAGGGAACGCCCCCGTTCAGGTTCACTTC
1621  CTGGATCCTTACTGCTCTGCCTCGGTGGCAGGAGCCCGGGAAGGAGATTATATTGTCTCC
1681  ATTCAGCTTGTGGATTGTAAGTGGCTGACGCTGAGTGAGGTTATGAAGCTGCTGAAGAGC
1741  TTTGGCGAGGACGAGATCGAGATGAAAGTCGTGAGCCTCCTGGACTCCACATCATCCATG
1801  CATAATAAGAGTGCCACATACTCCGTGGGAATGCAGAAAACGTACTCCATGATCTGCTTA
1861  GCCATTGATGATGACGACAAAACTGATAAAACCAAGAAAATCTCCAAGAAGCTTTCCTTC
1921  CTGAGTTGGGGCACCAACAAGAACAGACAGAAGTCAGCCAGCACCTTGTGCCTCCCATCG
1981  GTCGGGGCTGCACGGCCTCAGGTCAAGAAGAAGCTGCCCTCCCCTTTCAGCCTTCTCAAC
2041  TCAGACAGTTCTTGGTACTAATGTGAGGAAACAAACATGTTCAGGCCCCGAACATTTCCG
2101  GTGCTGACTCGGCCTTAAACGTTTGTGCCATAATGGAAAATATCTATCTATCTGTTCTCA
2161  AATCCTGTTTTTCTCATAGTGTAAACTCACATTTGATGTGTTTTTATGAAGGAAAGTAAC
2221  CAAGAAACCTCTAGGAATTAGTGAAAAAAGAACTTTTTTGAGGTGTGTTACTATACTGCT
2281  GTAAGTTATTTATTATATAAAGTATTGTAAATAGAATAGTGTTGAAGATATGAAATATGG
2341  CTATTTTTAATGGTGACAATTATGACTTTTAGTCACTATTAAATTGGGGTTACCTATATC
2401  AGTACAATTTGTAGTTGTTTCCAGGTTTGGCTAATAATCATTCCTTAACCTAGAATTCAG
2461  ATGATCCTGGAATTAAGGCAGGTCAGAGGACTGTAATGATAGAATTAAATTAGTGTCACT
2521  AAAAACTGTCCCAAAGTGCTGCTTCCTAATAGGAATTCATTAACCTAAAACAAGATGTTA
2581  CTATTATATCGATAGACTATGAATGCTATTTCTAGAAAAAGTCTAGTGCCAAATTTGTCT
2641  TATTAAATAAAAACAATGTAGGAGCAGCTTTTCTTCTAGTTTGATGTCATTTAAGAATTA
2701  CTAACACAGTGGCAGTGTTAGATGAAGATGCTGTCTACAAGGTAGATAATATACTGTTTG
2761  ATACTCAAAACATTTTTCATTTTGTTTAAAGTAGAAGTTACATAATTCTATATTTTAAGT
2821  CTTGGGTAAAAAAGTAGTTTTACATTTTATAAAGTAAAGATGTAAATGATTCAGGTTTAA
2881  AGCTCTATTTGACTTCCTTTTTTTGTTTGAGATAGCGTCTTGCTGTGTTGCCCAGGCTGG
2941  AGTGCAGTGGTGTGATCTCAGCTCAGTGCAACCTCCGCCCCCTGGGATCAAGCGATTCTC
3001  CTACCTCAGCCTCCCAAATAGCTGGGACTACAAGGTGCCCTCCAGCATGCCTGGCTGATT
3061  TTTGTATTTTTAGTTGAGGTGAGGTTTCACCATGTTGGCCAGGCGGGTTTCGAAATCCTG
3121  ACCTCAAATGATCCACCCACCTCAGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCAC
3181  CACAACCGTCCCACTATTTTACTTTTTAAAATGACATTCCTACTGATTGATTTTTATCTT
3241  GCTATAAGTTCGATGACACCGTGAATCTAATAAGGTTCACTGTTGACACAGTACAAGTTA
3301  CATAGCTAAAATACATAGCATTGAAGACTAATTTTAAGGATTGACAAGAGTTTATTTTCT
3361  ATTGTGCAATATCTTAAAGGAAGCAACCACCTTTGGGAAAGTGTATCTGCTGCTCCTAGG
3421  GCCATGCTTGTATACATATTTAAATAAACATATTCATTTACCCGAAAAAAAAAAAAAAAA
3481  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 14A

```
   1  M  T  D  A  L  L  P  A  A  P  Q  P  L  E  K  E  N  D  G  Y
   1  ATGACCGACGCGCTGTTGCCCGCGGCCCCCCAGCCGCTGGAGAAGGAGAACGACGGCTAC

  21  F  R  K  G  C  N  P  L  A  Q  T  G  R  S  K  L  Q  N  Q  R
  61  TTTCGGAAGGGCTGTAATCCCCTTGCACAAACCGGCCGGAGTAAATTGCAGAATCAAAGA

  41  A  A  L  N  Q  Q  I  L  K  A  V  R  M  R  T  G  A  E  N  L
 121  GCTGCTTTGAATCAGCAGATCCTGAAAGCCGTGCGGATGAGGACCGGAGCGGAAAACCTT

  61  L  K  V  A  T  N  S  K  V  R  E  Q  V  R  L  E  L  S  F  V
 181  CTGAAAGTGGCCACAAACTCAAAGGTGCGGGAGCAAGTGCGGCTGGAGCTGAGCTTCGTC

  81  N  S  D  L  Q  M  L  K  E  E  L  E  G  L  N  I  S  V  G  V
 241  AACTCAGACCTGCAGATGCTCAAGGAAGAGCTGGAGGGGCTGAACATCTCGGTGGGCGTC

 101  Y  Q  N  T  E  E  A  F  T  I  P  L  I  P  L  G  L  K  E  T
 301  TATCAGAACACAGAGGAGGCATTTACGATTCCCCTGATTCCTCTTGGCCTGAAGGAAACG

 121  K  D  V  D  F  A  V  V  L  K  D  F  I  L  E  H  Y  S  E  D
 361  AAAGACGTCGACTTTGCAGTCGTCCTCAAGGATTTTATCCTGGAACATTACAGTGAAGAT

 141  G  Y  L  Y  E  D  E  I  A  D  L  M  D  L  R  Q  A  C  R  T
 421  GGCTATTTATATGAAGATGAAATTGCAGATCTTATGGATCTGAGACAAGCTTGTCGGACG

 161  P  S  R  D  E  A  G  V  E  L  L  M  T  Y  F  I  Q  L  G  F
 481  CCTAGCCGGGATGAGGCCGGGGTGGAACTGCTGATGACATACTTCATCCAGCTGGGCTTT

 181  V  E  S  R  F  F  P  P  T  R  Q  M  G  L  L  F  T  W  Y  D
 541  GTCGAGAGTCGATTCTTCCCGCCCACACGGCAGATGGGACTCCTGTTCACCTGGTATGAC

 201  S  L  T  G  V  P  V  S  Q  Q  N  L  L  L  E  K  A  S  V  L
 601  TCTCTCACTGGGGTTCCGGTCAGCCAGCAGAACCTGCTGCTGGAGAAGGCCAGTGTCCTG

 221  F  N  T  G  A  L  Y  T  Q  I  G  T  R  C  D  R  Q  T  Q  A
 661  TTCAACACTGGGGCCCTCTACACCCAGATTGGGACCCGGTGCGATCGGCAGACGCAGGCT

 241  G  L  E  S  A  I  D  A  F  Q  R  A  A  G  V  L  N  Y  L  K
 721  GGGCTGGAGAGTGCCATAGATGCCTTTCAGAGAGCCGCAGGGGTTTTAAATTACCTGAAA

 261  D  T  F  T  H  T  P  S  Y  D  M  S  P  A  M  L  S  V  L  V
 781  GACACATTTACCCATACTCCAAGTTACGACATGAGCCCTGCCATGCTCAGCGTGCTCGTC

 281  K  M  M  L  A  Q  A  Q  E  S  V  F  E  K  I  S  L  P  G  I
 841  AAAATGATGCTTGCACAAGCCCAAGAAAGCGTGTTTGAGAAAATCAGCCTTCCTGGGATC

 301  R  N  E  F  F  M  L  V  K  V  A  Q  E  A  A  K  V  G  E  V
 901  CGGAATGAATTCTTCATGCTGGTGAAGGTGGCTCAGGAGGCTGCTAAGGTGGGAGAGGTC

 321  Y  Q  Q  L  H  A  A  M  S  Q  A  P  V  K  E  N  I  P  Y  S
 961  TACCAACAGCTACACGCAGCCATGAGCCAGGCGCCGGTGAAAGAGAACATCCCCTACTCC

 341  W  A  S  L  A  C  V  K  A  H  H  Y  A  A  L  A  H  Y  F  T
1021  TGGGCCAGCTTAGCCTGCGTGAAGGCCCACCACTACGCGGCCCTGGCCCACTACTTCACT

 361  A  I  L  L  I  D  H  Q  V  K  P  G  T  D  L  D  H  Q  E  K
1081  GCCATCCTCCTCATCGACCACCAGGTGAAGCCAGGCACGGATCTGGACCACCAGGAGAAG
```

FIGURE 14B

```
 381   C  L  S  Q  L  Y  D  H  M  P  E  G  L  T  P  L  A  T  L  K
1141   TGCCTGTCCCAGCTCTACGACCACATGCCAGAGGGGCTGACACCCTTGGCCACACTGAAG

 401   N  D  Q  Q  R  R  Q  L  G  K  S  H  L  R  R  A  M  A  H  H
1201   AATGATCAGCAGCGCCGACAGCTGGGGAAGTCCCACTTGCGCAGAGCCATGGCTCATCAC

 421   E  E  S  V  R  E  A  S  L  C  K  K  L  R  S  I  E  V  L  Q
1261   GAGGAGTCGGTGCGGGAGGCGAGCCTCTGCAAGAAGCTGCGGAGCATTGAGGTGCTACAG

 441   K  V  L  C  A  A  Q  E  R  S  R  L  T  Y  A  Q  H  Q  E  E
1321   AAGGTGCTGTGTGCCGCACAGGAACGCTCCCGGCTCACGTACGCCCAGCACCAGGAGGAG

 461   D  D  L  L  N  L  I  D  A  P  S  V  V  A  K  T  E  Q  E  V
1381   GATGACCTGCTGAACCTGATCGACGCCCCCAGTGTTGTTGCTAAAACTGAGCAAGAGGTT

 481   D  I  I  L  P  Q  F  S  K  L  T  V  T  D  F  F  Q  K  L  G
1441   GACATTATATTGCCCCAGTTCTCCAAGCTGACAGTCACGGACTTCTTCCAGAAGCTGGGC

 501   P  L  S  V  F  S  A  N  K  R  W  T  P  P  R  S  I  R  F  T
1501   CCCTTATCTGTGTTTTCGGCTAACAAGCGGTGGACGCCTCCTCGAAGCATCCGCTTCACT

 521   A  E  E  G  D  L  G  F  T  L  R  G  N  A  P  V  Q  V  H  F
1561   GCAGAAGAAGGGGACTTGGGGTTCACCTTGAGAGGGAACGCCCCCGTTCAGGTTCACTTC

 541   L  D  P  Y  C  S  A  S  V  A  G  A  R  E  G  D  Y  I  V  S
1621   CTGGATCCTTACTGCTCTGCCTCGGTGGCAGGAGCCCGGGAAGGAGATTATATTGTCTCC

 561   I  Q  L  V  D  C  K  W  L  T  L  S  E  V  M  K  L  L  K  S
1681   ATTCAGCTTGTGGATTGTAAGTGGCTGACGCTGAGTGAGGTTATGAAGCTGCTGAAGAGC

 581   F  G  E  D  E  I  E  M  K  V  V  S  L  L  D  S  T  S  S  M
1741   TTTGGCGAGGACGAGATCGAGATGAAAGTCGTGAGCCTCCTGGACTCCACATCATCCATG

 601   H  N  K  S  A  T  Y  S  V  G  M  Q  K  T  Y  S  M  I  C  L
1801   CATAATAAGAGTGCCACATACTCCGTGGGAATGCAGAAAACGTACTCCATGATCTGCTTA

 621   A  I  D  D  D  D  K  T  D  K  T  K  K  I  S  K  K  L  S  F
1861   GCCATTGATGATGACGACAAAACTGATAAAACCAAGAAAATCTCCAAGAAGCTTTCCTTC

 641   L  S  W  G  T  N  K  N  R  Q  K  S  A  S  T  L  C  L  P  S
1921   CTGAGTTGGGGCACCAACAAGAACAGACAGAAGTCAGCCAGCACCTTGTGCCTCCCATCG

 661   V  G  A  A  R  P  Q  V  K  K  K  L  P  S  P  F  S  L  L  N
1981   GTCGGGGCTGCACGGCCTCAGGTCAAGAAGAAGCTGCCCTCCCCTTTCAGCCTTCTCAAC

 681   S  D  S  S  W  Y  -
2041   TCAGACAGTTCTTGGTACTAATGTGAGGAAACAAACATGTTCAGGCCCCGAACATTTCCG
```

FIGURE 14B

```
TATA PROMOTER AND PUTATIVE TRANSCRIPTION START SITE
        AAAAAAAATAAATAAAAAGGCCGGGCGCGTTGGCCCGCGCcTGCAGCCCC

PSL 22_5'UTR
     1  TGCTACTTGGGAGGCTGAGGCTGGAGCATCGCTTGATCCTGGGAGGTCGAGGCTGCAAAG
    61  AGTCGAGATCGCAACACTGCTCTCCAGCCTGGGCGACAGAGCGAGGTCCCATCTCTTAAA
   121  AAAAAGAACTGTGCTCAAGGACATCTGCCGTGTCTGGGGCGCAAAACCCCTCCTGGTCCC
   181  CTCTCTCAGGGCAGTCCGCGAGCCCAGCGGATCCCACTCGTCTTTGCAGCGCGGACAGGG
   241  AATCGGCTGAGTTGATCCCATGCCAACAAGCCCGAGTAGTCCGGGCAAGGCGCTCGGCGG
   301  GGCAGTCAACGCTCCCTCCGCCATGGGCTCCCCTCTTGGGAAAAGCTTTTCCAAACCGCC
   361  GGGCCCAGGGCCCAGAGCTCCCGCCGCGCCCTCGACGTGGCGTCGAGTCTGGCCCCTTCC
   421  CCCGCGGCGCACGGGCTTCACCCAGGAGGGACGCGCCTGGATCCACGCCTTCCTCACTGA
   481  CTCCCCGGGCTCCAGGGCAGGGTGCAGGTCCACAGCCAGGGCTTCGCTGCGGCCCCTGAG
   541  ACCCCAGTGCCTTTCCTGCGCTCTCGCGGCACTCGCAAAGTTGAGTCAGCCACGACGCCC
   601  ACAGACAACCCCGAGGCGCCGCGCCCAGGGCGCAGCTCTCCGGGTGACGAGCGCCTCAAG
   661  GGGCGCGGGTTCGGGGCCCGCGACGGGGCGGGGCGCGTCTCCAGGGCTCCAGTGCTCGGC
   721  CTCAGGCGGGGCTAGAAGGGCCGCGGGACGGGGTGGGAGTGGAGGGGCGGGGAAGGGCGG
   781  GGACAGGGGCGGGGCCGCACGTCCTCTCGGGCCAGCCTCAGCCGCCGCGCCTCAGTCCGC
   841  CGTCCGCCCTCCGCGCCCGCGCCGCTAGC

EXON_1 69bp
     1  ATGACCGACGCGCTGTTGCCCGCGGCCCCCCAGCCGCTGGAGAAGGAGAACGACGGCTAC
    61  TTTCGGAAG

EXON_2 117bp
     1  GGCTGTAATCCCCTTGCACAAACCGGCCGGAGTAAATTGCAGAATCAAAGAGCTGCTTTG
    61  AATCAGCAGATCCTGAAAGCCGTGCGGATGAGGACCGGAGCGGAAAACCTTCTGAAA

EXON_3 129bp
     1  GTGGCCACAAACTCAAAGGTGCGGGAGCAAGTGCGGCTGGAGCTGAGCTTCGTCAACTCA
    61  GACCTGCAGATGCTCAAGGAAGAGCTGGAGGGGCTGAACATCTCGGTGGGCGTCTATCAG
   121  AACACAGAG

EXON_4 75bp
     1  GAGGCATTTACGATTCCCCTGATTCCTCTTGGCCTGAAGGAAACGAAAGACGTCGACTTT
    61  GCAGTCGTCCTCAAG

EXON_5 79bp
     1  GATTTTATCCTGGAACATTACAGTGAAGATGGCTATTTATATGAAGATGAAATTGCAGAT
    61  CTTATGGATCTGAGACAAG

EXON_6 124bp
     1  CTTGTCGGACGCCTAGCCGGGATGAGGCCGGGGTGGAACTGCTGATGACATACTTCATCC
    61  AGCTGGGCTTTGTCGAGAGTCGATTCTTCCCGCCCACACGGCAGATGGGACTCCTGTTCA
   121  CCTG

EXON_7 167bp
     1  GTATGACTCTCTCACTGGGGTTCCGGTCAGCCAGCAGAACCTGCTGCTGGAGAAGGCCAG
    61  TGTCCTGTTCAACACTGGGGCCCTCTACACCCAGATTGGGACCCGGTGCGATCGGCAGAC
   121  GCAGGCTGGGCTGGAGAGTGCCATAGATGCCTTTCAGAGAGCCGCAG

EXON_8 188bp
     1  GGGTTTTAAATTACCTGAAAGACACATTTACCCATACTCCAAGTTACGACATGAGCCCTG
    61  CCATGCTCAGCGTGCTCGTCAAAATGATGCTTGCACAAGCCCAAGAAAGCGTGTTTGAGA
   121  AAATCAGCCTTCCTGGGATCCGGAATGAATTCTTCATGCTGGTGAAGGTGGCTCAGGAGG
   181  CTGCTAAG
EXON_9 156bp
     1  GTGGGAGAGGTCTACCAACAGCTACACGCAGCCATGAGCCAGGCGCCGGTGAAAGAGAAC
```

FIGURE 14C

```
    61  ATCCCCTACTCCTGGGCCAGCTTAGCCTGCGTGAAGGCCCACCACTACGCGGCCCTGGCC
   121  CACTACTTCACTGCCATCCTCCTCATCGACCACCAG

EXON_10 120bp
     1  GTGAAGCCAGGCACGGATCTGGACCACCAGGAGAAGTGCCTGTCCCAGCTCTACGACCAC
    61  ATGCCAGAGGGGCTGACACCCTTGGCCACACTGAAGAATGATCAGCAGCGCCGACAGCTG

EXON_11 196bp
     1  GGGAAGTCCCACTTGCGCAGAGCCATGGCTCATCACGAGGAGTCGGTGCGGGAGGCGAGC
    61  CTCTGCAAGAAGCTGCGGAGCATTGAGGTGCTACAGAAGGTGCTGTGTGCCGCACAGGAA
   121  CGCTCCCGGCTCACGTACGCCCAGCACCAGGAGGAGGATGACCTGCTGAACCTGATCGAC
   181  GCCCCCAGTGTTGTTG

EXON_12 77bp
     1  CTAAAACTGAGCAAGAGGTTGACATTATATTGCCCCAGTTCTCCAAGCTGACAGTCACGG
    61  ACTTCTTCCAGAAGCTG

EXON_13 147bp
     1  GGCCCCTTATCTGTGTTTTCGGCTAACAAGCGGTGGACGCCTCCTCGAAGCATCCGCTTC
    61  ACTGCAGAAGAAGGGGACTTGGGGTTCACCTTGAGAGGGAACGCCCCCGTTCAGGTTCAC
   121  TTCCTGGATCCTTACTGCTCTGCCTCG

EXON_14 156bp
     1  GTGGCAGGAGCCCGGGAAGGAGATTATATTGTCTCCATTCAGCTTGTGGATTGTAAGTGG
    61  CTGACGCTGAGTGAGGTTATGAAGCTGCTGAAGAGCTTTGGCGAGGACGAGATCGAGATG
   121  AAAGTCGTGAGCCTCCTGGACTCCACATCATCCATG

EXON_15 +3'UTR 1664bp+polyA tract
     1  CATAATAAGAGTGCCACATACTCCGTGGGAATGCAGAAAACGTACTCCATGATCTGCTTA
    61  GCCATTGATGATGACGACAAAACTGATAAAACCAAGAAAATCTCCAAGAAGCTTTCCTTC
   121  CTGAGTTGGGGCACCAACAAGAACAGACAGAAGTCAGCCAGCACCTTGTGCCTCCCATCG
   181  GTCGGGGCTGCACGGCCTCAGGTCAAGAAGAAGCTGCCCTCCCCTTTCAGCCTTCTCAAC
   241  TCAGACAGTTCTTGGTACTAATGTGAGGAAACAAACATGTTCAGGCCCCGAACATTTCCG
   301  GTGCTGACTCGGCCTTAAACGTTTGTGCCATAATGGAAAATATCTATCTATCTGTTCTCA
   361  AATCCTGTTTTTCTCATAGTGTAAACTCACATTTGATGTGTTTTTATGAAGGAAAGTAAC
   421  CAAGAAACCTCTAGGAATTAGTGAAAAAAGAACTTTTTTGAGGTGTGTTACTATACTGCT
   481  GTAAGTTATTTATTATATAAAGTATTGTAAATAGAATAGTGTTGAAGATATGAAATATGG
   541  CTATTTTTAATGGTGACAATTATGACTTTTAGTCACTATTAAATTGGGGTTACCTATATC
   601  AGTACAATTTGTAGTTGTTTCCAGGTTTGGCTAATAATCATTCCTTAACCTAGAATTCAG
   661  ATGATCCTGGAATTAAGGCAGGTCAGAGGACTGTAATGATAGAATTAAATTAGTGTCACT
   721  AAAAACTGTCCCAAAGTGCTGCTTCCTAATAGGAATTCATTAACCTAAAACAAGATGTTA
   781  CTATTATATCGATAGACTATGAATGCTATTTCTAGAAAAAGTCTAGTGCCAAATTTGTCT
   841  TATTAAATAAAAACAATGTAGGAGCAGCTTTTCTTCTAGTTTGATGTCATTTAAGAATTA
   901  CTAACACAGTGGCAGTGTTAGATGAAGATGCTGTCTACAAGGTAGATAATATACTGTTTG
   961  ATACTCAAAACATTTTTCATTTTGTTTAAAGTAGAAGTTACATAATTCTATATTTTAAGT
  1021  CTTGGGTAAAAAAGTAGTTTTACATTTTATAAAGTAAAGATGTAAATGATTCAGGTTTAA
  1081  AGCTCTATTTGACTTCCTTTTTTTGTTTGAGATAGCGTCTTGCTGTGTTGCCCAGGCTGG
  1141  AGTGCAGTGGTGTGATCTCAGCTCAGTGCAACCTCCGCCCCCTGGGATCAAGCGATTCTC
  1201  CTACCTCAGCCTCCCAAATAGCTGGGACTACAAGGTGCCCTCCAGCATGCCTGGCTGATT
  1261  TTTGTATTTTTAGTTGAGGTGAGGTTTCACCATGTTGGCCAGGCGGGTTTCGAAATCCTG
  1321  ACCTCAAATGATCCACCCACCTCAGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCAC
  1381  CACAACCGTCCCACTATTTTACTTTTTAAAATGACATTCCTACTGATTGATTTTTATCTT
  1441  GCTATAAGTTCGATGACACCGTGAATCTAATAAGGTTCACTGTTGACACAGTACAAGTTA
  1501  CATAGCTAAAATACATAGCATTGAAGACTAATTTTAAGGATTGACAAGAGTTTATTTTCT
  1561  ATTGTGCAATATCTTAAAGGAAGCAACCACCTTTGGGAAAGTGTATCTGCTGCTCCTAGG
  1621  GCCATGCTTGTATACATATTTaaataaACATATTCATTTACCCGAAAAAAAAAAAAAAAA
  1681  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 14C

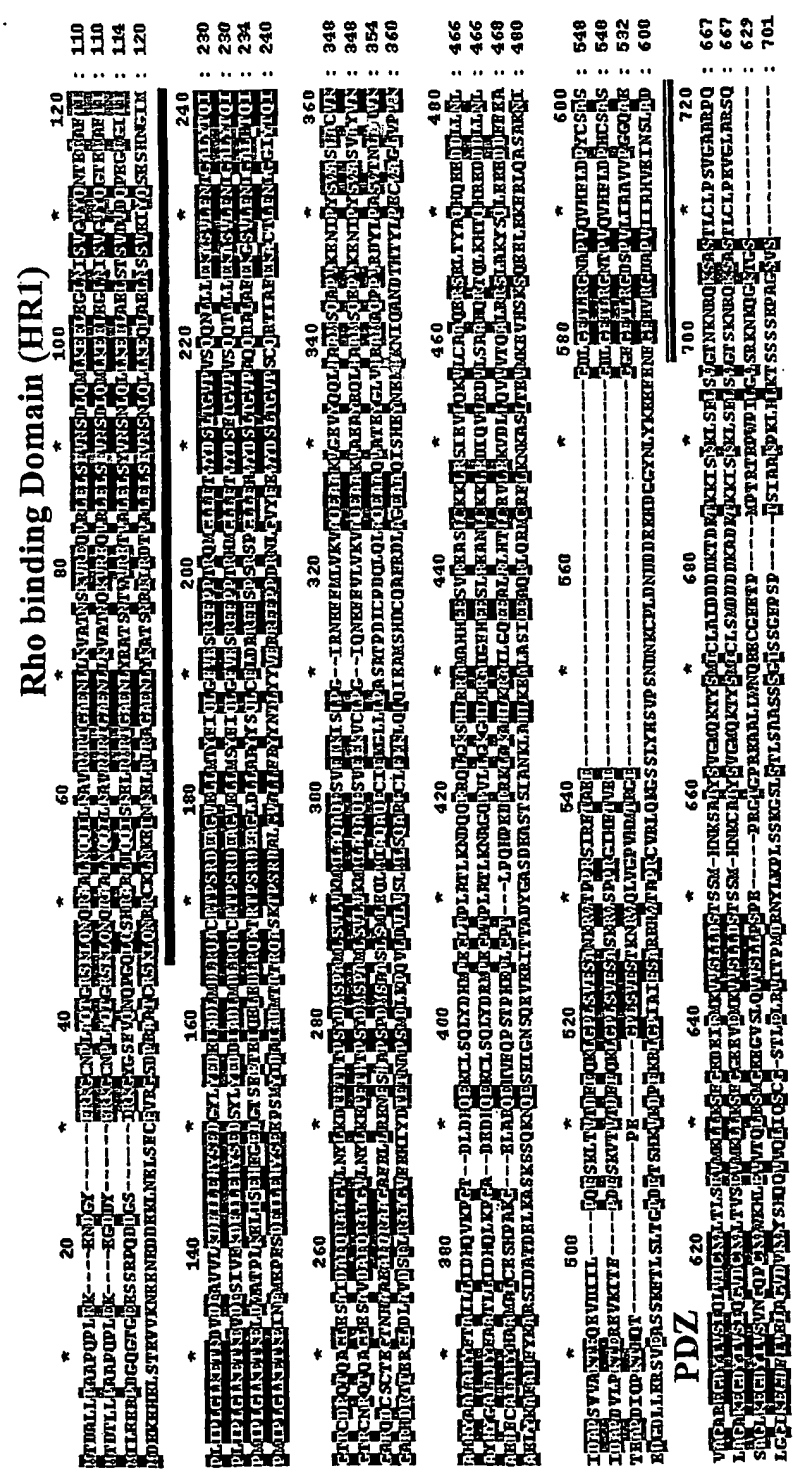


Figure 15

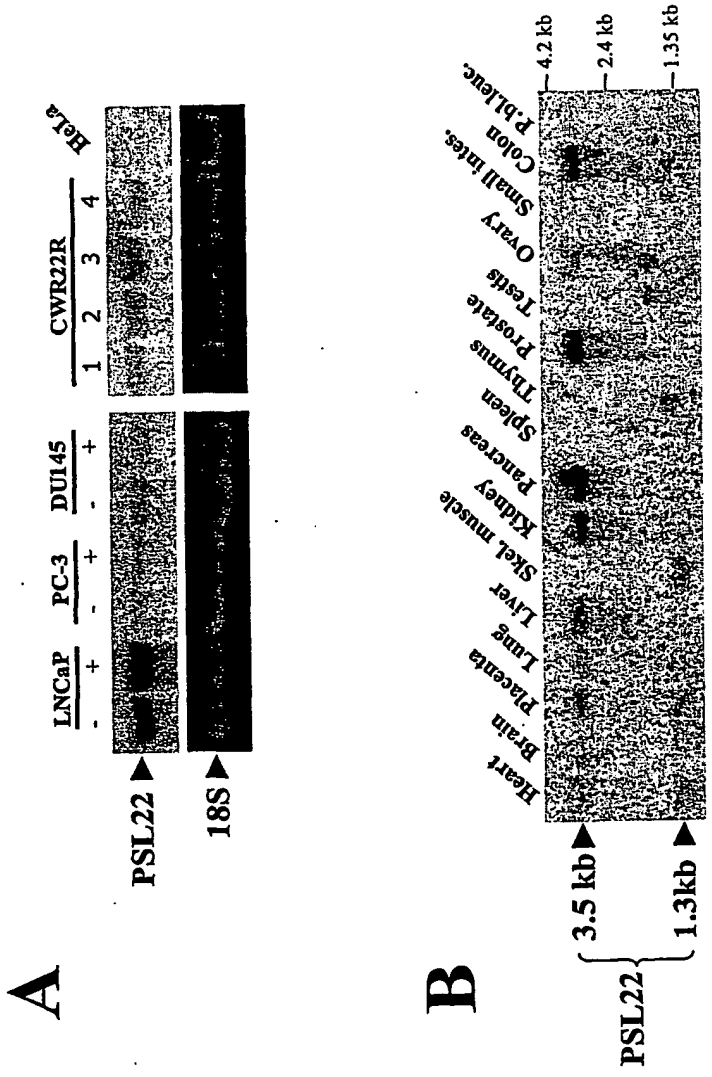

A
LNCaP
PC-3
DU145
CWR22R
HeLa
PSL22
18S
B
Heart
Brain
Placenta
Lung
Liver
Skel. muscle
Kidney
Pancreas
Spleen
Thymus
Prostate
Testis
Ovary
Small Intes.
Colon
P. bl.leuc.
4.2 kb
2.4 kb
1.35 kb
3.5 kb
1.3kb
PSL22

Figure 16